

EDUCACIÓN, CREATIVIDAD E INTELIGENCIA ARTIFICIAL: NUEVOS HORIZONTES PARA EL APRENDIZAJE. ACTAS DEL VIII CONGRESO INTERNACIONAL SOBRE APRENDIZAJE, INNOVACIÓN Y COOPERACIÓN, CINAIC 2025

María Luisa Sein-Echaluce Lacleta, Ángel Fidalgo Blanco y Francisco José García Peñalvo (coords.)

1ª Edición. Zaragoza, 2025

Edita: Servicio de Publicaciones. Universidad de Zaragoza.



Servicio de
Publicaciones
Universidad Zaragoza

EBOOK ISBN 978-84-10169-60-9

DOI 10.26754/uz.978-84-10169-60-9



Esta obra se encuentra bajo una licencia Creative Commons Reconocimiento – NoComercial (ccBY-NC). Ver descripción de esta licencia en <https://creativecommons.org/licenses/by-nc-nd/4.0/>

Referencia a esta obra:

Sein-Echaluce Lacleta, M.L., Fidalgo Blanco, A. & García-Peñalvo, F.J. (coords.) (2025). *Educación, Creatividad e Inteligencia Artificial: nuevos horizontes para el Aprendizaje. Actas del VIII Congreso Internacional sobre Aprendizaje, Innovación y Cooperación. CINAIC 2025 (11-13 de Junio de 2025, Madrid, España)*. Zaragoza. Servicio de Publicaciones Universidad de Zaragoza. DOI 10.26754/uz.978-84-10169-60-9

Análisis del Uso de Modelos Generativos de Inteligencia Artificial en el Aprendizaje de la Programación de Ordenadores

Analysis of the Use of Generative Artificial Intelligence Models in Computer Programming Learning

Laura de León-Torres, Ascensión Lopez-Vargas, Ángel García-Beltrán, Javier Rodríguez-Vidal, Pablo Manuel Vigarra Gallego

l.deleon@alumnos.upm.es, a.lvargas@upm.es, agarcia@etsii.upm.es, javier.rodriguez.vidal@upm.es, pm.vigara@upm.es

Departamento de Automática, Ingeniería Eléctrica y Electrónica e Informática Industrial
Universidad Politécnica de Madrid
Madrid, España

Resumen- El uso de herramientas de Inteligencia Artificial (IA) generativas ha aumentado notablemente en el ámbito educativo, entre otras, en la docencia de la programación. Este trabajo tiene como objetivo analizar el uso de estas herramientas en el aprendizaje de la programación de ordenadores mediante (a) un estudio del estado del arte, (b) una encuesta a los estudiantes, (c) una comprobación experimental de la eficiencia de una IA generativa (IAG) y (d) una actividad práctica con alumnos. La encuesta reveló que el 97% de los encuestados ha usado alguna IA, pero solo el 63% para programar. La comprobación experimental detectó errores recurrentes de estas herramientas generando código y dificultades en la resolución de ciertos tipos de ejercicios. En la actividad práctica, solo el 54% logró corregir un código con la ayuda de una IAG. Además, algunos, debido a su confianza en estas tecnologías, consideraron que algunos códigos eran correctos sin verificarlo.

Palabras clave: *Inteligencia Artificial Generativa, ChatGPT, Código Fuente, Programación, Docencia, Evaluación*

Abstract- The use of generative Artificial Intelligence (AI) tools has seen a notable rise in the educational field, among others, in the teaching of programming. The aim of this work is to analyse the use of these tools in computer programming learning through (a) a state of the art review, (b) a survey conducted with students, (c) an experimental verification of the efficiency of a generative AI and (d) a practical activity with students. The survey shows that 97% of the students surveyed have used some AI system, but only 63% for programming. The experimental verification detected recurrent errors made by these tools generating code and difficulties with certain types of exercises. In the practical activity, only 54% managed to correct a code with the help of a generative AI. In addition, some students, due to their overconfidence in these technologies, assumed that certain codes were correct without checking it correctly.

Keywords: *Generative Artificial Intelligence, ChatGPT, Source Code, Programming, Teaching, Evaluation*

1. INTRODUCCIÓN

En los últimos años, el uso de la Inteligencia Artificial (IA) ha experimentado un auge significativo, especialmente desde el lanzamiento de tecnologías como ChatGPT (OpenAI, s.f.) en noviembre de 2022. Este evento marcó un punto de inflexión que supuso la democratización de esta tecnología a través de la expansión de su acceso y aplicación entre usuarios no expertos.

Con la introducción de las IAs Generativas (IAG), ya no solo se procesa información, sino que ahora también se puede generar contenido, obtener razonamientos en diferentes contextos o resolver problemas.

El uso de las herramientas de IAG ha llegado también al ámbito educativo, principalmente a institutos y universidades, donde no han pasado desapercibidas. El alumnado, habituado a una constante evolución tecnológica, ha incorporado con rapidez estas plataformas en sus procesos de aprendizaje. ChatGPT es utilizado por los alumnos para la creación de texto y de ideas, la resolución de ejercicios y así como la generación y depuración de código fuente en el aprendizaje de la programación de ordenadores.

El empleo de estas tecnologías en educación ha provocado una serie de interrogantes y preocupaciones, surgiendo preguntas como: ¿es beneficioso el uso de estas herramientas para el aprendizaje de los alumnos?, ¿hasta qué punto es ético su empleo?, ¿podría considerarse plagio?, ¿puede desencadenar desigualdades entre los estudiantes?, ¿puede provocar que adquieran conceptos erróneos?

En este trabajo de investigación, se realizó un análisis del estado del arte centrado en la aplicación de las IAG para la creación de código fuente. Como resultado de este estudio se observó que la adopción de estos nuevos sistemas implica tanto beneficios como limitaciones, evidenciando la necesidad de un manejo equilibrado y supervisado en entornos educativos. Además, entre todas las plataformas de IAG, ChatGPT fue identificada como la más relevante en la literatura, siendo la IAG más utilizada en los estudios e investigaciones analizados.

ChatGPT destaca por su capacidad para generar soluciones, proponer ejemplos de casos concretos, responder dudas y corregir errores. En 2023, Sánchez Allende (Sánchez Allende, 2023) propuso la adopción de ChatGPT como tutor en asignaturas de programación, aprovechando su capacidad para facilitar un aprendizaje personalizado. De manera análoga, la literatura revisada plantea incertidumbres sobre si el empleo de IAGs puede tener un impacto negativo en el proceso de aprendizaje y la adquisición de conocimientos entre los alumnos. En primer lugar, esta aplicación puede generar soluciones con errores, proporcionando información incorrecta o inexacta. Por ejemplo, se ha observado que ChatGPT comete

errores al realizar cálculos matemáticos debido a que no está entrenado para ello y además carece de mecanismos para verificar la precisión de su salida (Anderson et al., 2023). La generación de soluciones incorrectas puede provocar que los estudiantes adopten conceptos erróneos como correctos, lo que afecta gravemente a su aprendizaje. Además, el profesorado señaló que el empleo excesivo de esta tecnología podría generar dependencia en el alumnado, lo cual dificultaría la adquisición y el desarrollo de habilidades esenciales, como la resolución de problemas, la creatividad o el pensamiento crítico.

En el estado del arte, se han documentado diversos estudios sobre la aplicación de IAG en el aprendizaje de la programación. Estas investigaciones abarcan encuestas dirigidas a estudiantes y docentes, evaluaciones de las herramientas por parte de programadores expertos y actividades prácticas con el alumnado. En 2022, se llevó a cabo un estudio con distintos sistemas basados en IAG aplicados a la creación de código fuente entre programadores (Vaithilingam et al., 2022). Se observó que los participantes que utilizaron Intellisense (Microsoft, s.f.) completaron con más frecuencia las tareas. Por su parte, los tiempos medios fueron inferiores usando Copilot (GitHub, s.f.). En la encuesta final, los participantes manifestaron su preferencia por Copilot. Los investigadores concluyeron que, aunque los programadores muestran un alto grado de aceptación hacia esta herramienta, es necesario ser cauteloso con la subestimación del tiempo requerido para corregir errores en el código propuesto, ya que la depuración se complica cuando uno mismo no ha escrito el código.

A principios de 2023, se realizó una encuesta a 20 profesores de programación de nueve países diferentes, con el objetivo de identificar cómo planeaban los profesores universitarios adaptar sus cursos tras la llegada de ChatGPT para garantizar el correcto aprendizaje entre los estudiantes (Lau & Guo, 2023). Tras la realización de distintas entrevistas, los investigadores identificaron dos enfoques en la respuesta del profesorado ante el uso de IAGs. A corto plazo, los docentes manifestaron preocupación por la posible manipulación indebida de estas tecnologías en la realización de exámenes y tareas académicas. En consecuencia, implementaron medidas como la modificación de los criterios de evaluación, la prohibición de estas tecnologías en determinadas actividades y la orientación a los estudiantes sobre sus capacidades y limitaciones. A largo plazo, se identificaron dos posturas principales entre los docentes. Por un lado, algunos rechazan la incorporación de estas plataformas en el proceso educativo, argumentando posibles efectos negativos en el desarrollo del pensamiento crítico y la resolución de problemas. Por otro lado, un grupo de docentes aboga por la integración progresiva de las IAG en la enseñanza, con el objetivo de aprovechar su potencial para mejorar el aprendizaje y la adquisición de habilidades digitales.

En relación con la evaluación de las herramientas, profesores de la Universidad de La Laguna (Sande & López Ramos, 2023) llevaron a cabo una serie de experiencias con ChatGPT, detectando que, en algunos casos, esta IAG no proporcionaba una solución aceptable de inmediato y que, para obtener un resultado correcto, hubo que insistir en la petición repitiendo algunas partes del enunciado del problema. También constataron que cambiar la forma de la solicitud o el idioma influye significativamente en la solución generada. Como conclusión, determinaron que la capacidad de ChatGPT para

realizar ejercicios sencillos de programación es satisfactoria, logrando resolver la mayoría de los ejercicios.

En relación con las actividades experimentales con alumnos, en abril de 2023 se realizó un estudio con el objetivo de medir la efectividad de ChatGPT en el proceso de aprendizaje (Qureshi, 2023). Para ello, se llevó a cabo un experimento con 24 estudiantes que fueron divididos en dos grupos equilibrados. A ambos grupos se les proporcionó los mismos ejercicios y se calificaron en función del número de casos exitosos. En los resultados de esta experimentación se observó que el grupo que empleaba ChatGPT, consiguió una puntuación significativamente mayor en todos los ejercicios. Sin embargo, se constata que, ningún equipo de este grupo, en los ejercicios de mayor dificultad obtuvo la puntuación máxima. Los investigadores concluyeron que ChatGPT pierde su efectividad cuanto mayor es la dificultad del problema y que por ello, los docentes tienen que incentivar el empleo adecuado de estos recursos. Por otra parte, en la Universidad de Colombia se realizó una experimentación que consistió en un estudio con un grupo de 51 universitarios, de los cuales el 96% ya conocía ChatGPT (Terán, 2023). Al analizar los resultados, el profesor constató que el número de ejercicios resueltos de manera correcta aumentó un 61% al emplear de forma ética ChatGPT. Además, observó una mejora significativa en la comprensión por parte de los alumnos y un aumento de la motivación, especialmente entre aquellos con más dificultades al inicio de la asignatura.

2. CONTEXTO Y DESCRIPCIÓN

El objetivo principal de este trabajo es evaluar la eficiencia de las herramientas de IAG generando código fuente, y su influencia en la docencia de la asignatura Fundamentos de Programación de la Escuela Técnica Superior de Ingenieros Industriales (ETSII) de la Universidad Politécnica de Madrid (UPM).

Con la finalidad de alcanzar el objetivo principal, se ha desarrollado una metodología formada por (A) una encuesta sobre el uso de las aplicaciones de IAG entre alumnos, (B) comprobación de la resolución de ejercicios propuestos por una IAG y (C) una actividad experimental para evaluar el uso de ChatGPT por parte de los alumnos en las prácticas de la asignatura de Fundamentos de Programación.

A. Encuesta sobre el uso de herramientas IAG entre los alumnos

Se realiza una encuesta a 65 alumnos universitarios de la ETSII-UPM para conocer mejor como emplean las herramientas de IAG los estudiantes y, además, conocer sus opiniones y preferencias. Esta encuesta se creó con el formato de un formulario de Google Forms y se compartió a través de WhatsApp por grupos de alumnos de la escuela.

La encuesta constaba de once preguntas que fueron distribuidas en cuatro secciones distintas. La primera sección clasificaba a los encuestados según su grado, curso y uso de herramientas de IAG. La segunda y tercera secciones exploraban la utilización de estos recursos por parte de los estudiantes y su nivel de satisfacción, tanto de forma general, en la segunda, como para la generación de código, en la tercera. Finalmente, la última sección recogió la percepción de aquellos que no han utilizado ningún sistema de IAG durante su carrera académica.

B. Comprobación de la resolución de ejercicios propuestos por ChatGPT

Para determinar si ChatGPT puede ayudar a los alumnos con su aprendizaje de los fundamentos de programación, es esencial conocer primero sus puntos fuertes y débiles en este ámbito. En este estudio se analizó cómo ChatGPT es capaz de resolver ejercicios numéricos y problemas de programación de la asignatura Fundamentos de Programación de diferentes grados impartidos en la ETSII-UPM. En esta asignatura se programa en el estándar C89/C90.

ChatGPT es una herramienta de IAG desarrollada en noviembre de 2022 por OpenAI y desde entonces han aparecido varias versiones. Sin embargo, la comprobación experimental de su eficiencia se ha realizado con la versión ChatGPT 3.5, ya que, en el momento de realizar la evaluación, era la versión gratuita disponible más actual y usada por el alumnado.

Debido a la existencia de una gran colección de exámenes y sus respectivas soluciones, en esta evaluación experta se le pide a ChatGPT que resuelva ejercicios de exámenes de los últimos años. Para cada ejercicio, se abre un chat individual con ChatGPT y se le entrega el enunciado correspondiente. En total se ha realizado la evaluación experta de ChatGPT en la resolución de 99 ejercicios de exámenes de años anteriores. Los indicadores que se usaron para esta evaluación fueron (1) el número de problemas que esta IAG consigue resolver y (2) el número de intervenciones que se necesitan para alcanzar la resolución correcta de cada ejercicio.

C. Uso de ChatGPT por parte de los alumnos en las prácticas de Fundamentos de Programación

Para poder obtener una información más realista sobre cómo los estudiantes manejan ChatGPT, se realizó una práctica de laboratorio *ad hoc*.

Se decidió realizar esta experimentación durante la última sesión de prácticas de la asignatura Fundamentos de Programación del Grado en Ingeniería en Tecnologías Industriales (GITI) de la ETSII UPM del curso académico 2023-2024. La temática de esta práctica fue el manejo de Archivos en C. Participaron 341 alumnos de nuevo ingreso matriculados en la asignatura, repartidos en 18 grupos de prácticas.

Durante esta experimentación, se les proporcionó a los estudiantes participantes en el experimento un enunciado junto a un código generado como respuesta por ChatGPT que contenía errores y no cumplía con los objetivos del problema. Se les pidió que, con la única ayuda de ChatGPT, corrigieran el código en menos de diez minutos. Al finalizar debían responder un cuestionario indicando si: (a) habían conseguido corregir el código, haciendo que compilara y que no contara con errores lógicos y (b) el número de intentos que habían necesitado para conseguirlo.

3. RESULTADOS

A. Resultados de la encuesta realizada sobre el uso de herramientas IAG entre alumnos

Los resultados de la encuesta mostraron que una gran proporción de los participantes, el 97% de ellos, conocen y usan las herramientas de IAG de manera habitual. Solo el 3% de los encuestados indicaron que no las habían utilizado nunca. Por

otro lado, la proporción de encuestados que habían usado alguna de estas plataformas para la tarea concreta de generación de código fuente era menor: el 63% de los alumnos encuestados. Ambos datos muestran la gran acogida que han tenido estas tecnologías entre los estudiantes de la universidad.

Además, se observó que la satisfacción con estos sistemas, en la mayoría de los casos, es intermedia, tanto en su aplicación general, con una satisfacción media de 3.39 sobre 5, como para la escritura de código fuente, con una satisfacción media de 3.51 sobre 5, ya que estas herramientas cometen fallos y pueden proporcionar respuestas erróneas. La versión gratuita de ChatGPT resalta como la IAG más empleada entre los estudiantes, tanto para tareas diversas como para tareas de programación.

B. Análisis de la eficacia de ChatGPT en la resolución de ejercicios de programación realizada por un experto

En la realización de la evaluación de la eficiencia de la versión gratuita de ChatGPT 3.5 generando código fuente en el estándar C89/C90, se detectó en primer lugar la existencia de lo que se ha denominado “errores típicos”, fallos que esta herramienta siempre comete al resolver ejercicios de programación en C. Estos errores típicos son: la forma de comentar, la declaración de variables en bucles *for* y el uso de *const char **. Debido a la aparición de manera recurrente de ambos errores, se decidió no incluir las intervenciones dedicadas a su resolución en el análisis.

Como resultado de esta evaluación, se evidencia que esta herramienta puede ser un recurso útil para resolver ejercicios básicos de codificación numérica en binario. Resuelve correctamente el 67% de los ejercicios de codificación. Además, se detecta que muestra un mejor rendimiento en los ejercicios con tareas más sencillas, como los relacionados con frecuencias, número de bits y resultados de expresiones. Sin embargo, para los ejercicios con tareas más complejas, como conversiones de bytes, el modelo requiere intervenciones para alcanzar una respuesta correcta. Esto refleja que, en los casos donde las conversiones o cálculos implican varios pasos intermedios, ChatGPT encuentra dificultades para seguir el proceso de manera precisa sin ser guiado por el usuario.

Además, se constata que ChatGPT tiene dificultades con los problemas que involucran números en coma flotante, en particular cuando se utilizan formatos más avanzados como la mantisa y el exponente en complemento a dos. Estos fallos indican que, aunque el modelo puede resolver muchos ejercicios de codificación más sencillos, su capacidad para manejar formatos numéricos complejos y que requieren una comprensión profunda de ciertos conceptos es muy limitada.

En segundo lugar, se analizan ejercicios de programación y se detecta que, aunque ChatGPT es capaz de generar soluciones correctas el 81% de las veces, en ocasiones comete errores no tanto de programación (excluyendo los errores típicos ya mencionados), sino de comprensión del enunciado. Estos errores conceptuales son corregibles mediante intervenciones, lo que destaca la importancia de una revisión cuidadosa por parte del usuario para asegurar la precisión en las soluciones propuestas.

Por otro lado, los ejercicios que necesitan más intervenciones por parte del usuario son aquellos relacionados con la manipulación de archivos. Esto se debe a que la complejidad de este tema es mayor. Desde la generación de archivos, lectura y

escritura en archivos hasta la manipulación de datos almacenados externamente. Además de que la interpretación precisa del enunciado juega un papel importante, ya que cualquier ambigüedad o mala interpretación por parte del modelo puede llevar a resultados incorrectos. Además, se tiende a combinar estos ejercicios con conceptos de otros temas de la asignatura, lo que eleva su dificultad.

C. Resultados de la experimentación con ChatGPT de los alumnos en las prácticas de Fundamentos de Programación

En esta experimentación se ha evidenciado que no todos los alumnos que han participado tienen las mismas habilidades con ChatGPT, ya que solo el 54% consiguió alcanzar el objetivo, mientras que el 46% restante no. Existiendo incluso estudiantes (1.8%) que no quisieron participar en la experiencia y prefirieron no emplear esta IAG.

Además, se ha detectado un fenómeno preocupante entre algunos participantes: la tendencia de considerar que el código era correcto y funcionaba simplemente porque no aparecían errores de compilación, sin comprobar si ejecutaba adecuadamente y estaba libre de errores lógicos.

En cuanto al número de preguntas que necesitaron realizar a ChatGPT durante la experiencia, este varía considerablemente. La media de preguntas de aquellos que indican que han logrado el objetivo es de 4.6 mientras que la de los que no lo consiguen es significativamente superior, de 9.4. Un alumno llegó a realizar 40 preguntas en los diez minutos que duraba el experimento. Esto es un indicio de que algunos estudiantes no saben emplear estos recursos de manera eficiente.

En resumen, los resultados de la experimentación práctica con estudiantes evidenciaron la necesidad de una formación específica para que el alumnado aprenda a utilizar este tipo de herramientas sin que comprometan su correcto aprendizaje.

4. CONCLUSIONES

Este estudio analiza la capacidad de las herramientas de IAG generando código fuente, centrándose en la evaluación de la versión gratuita 3.5 de ChatGPT en la docencia de la asignatura Fundamentos de Programación en la ETSII UPM.

Los resultados de la encuesta a los estudiantes indican una alta penetración de las plataformas de IAG entre los encuestados, con un 97% reportando su uso y un 63% empleándolas para programar. Además, la versión gratuita de ChatGPT destaca como la herramienta más utilizada por los estudiantes.

La evaluación experta del desempeño de ChatGPT resolviendo ejercicios de programación, permitió identificar áreas en las que ChatGPT destaca, como la resolución de ejercicios básicos relacionados con frecuencias y cálculos de bits. También se identificaron ciertas limitaciones en tareas complejas, como la manipulación de números en coma flotante y representaciones avanzadas. Los principales problemas detectados incluyeron errores de interpretación del enunciado.

En la experimentación con los alumnos, se observó que, aunque ChatGPT resulta útil para detectar y corregir errores básicos en el código, su utilidad depende en gran medida de la habilidad del estudiante, es decir, de su conocimiento y de las instrucciones que le aporta para guiarla y validar sus respuestas. Por lo que se concluye que existe la necesidad de formar sobre

el uso de estas herramientas para que los estudiantes aprendan a usarlas sin que perjudique su aprendizaje.

Es importante mencionar que el estudio presenta una limitación derivada del uso de la versión ChatGPT 3.5, ya que en el momento de su realización era la versión gratuita más actual y la empleada por el alumnado. Sin embargo, estas herramientas evolucionan a gran velocidad, por lo que los resultados obtenidos podrían quedar invalidados frente a las capacidades de las versiones más recientes. Por ello, sería interesante ampliar el análisis incorporando estas versiones actualizadas.

AGRADECIMIENTOS

Este trabajo ha sido financiado por la Universidad Politécnica de Madrid a través de la Convocatoria 2023-2024 de Proyectos de Innovación Educativa (Uso de Inteligencias Artificiales Generativas en el Aprendizaje de la Programación de Ordenadores, código de proyecto IE24.0504).

REFERENCIAS

- Anderson, N., McGowan, A., Galway, L., Hanna, P., Collins, M., & Cutting, D. (2023). Implementing Generative AI and Large Language Models in Education.
- GitHub. (s.f.). GitHub Copilot. <<https://github.com/features/copilot>>. Visitado por última vez 20.02.2025.
- Lau, S., & Guo, P. (2023). From “Ban It Till We Understand It” to “Resistance is Futile”: How University Programming Instructors Plan to Adapt as More Students Use AI Code Generation and Explanation Tools such as ChatGPT and GitHub Copilot. ICER '23 V1: Proceedings of the 2023 ACM Conference on International Computing Education Research V.1, Chicago, IL, USA, August 2023.
- Microsoft. (s.f.). IntelliSense in Visual Studio Code. <<https://code.visualstudio.com/docs/editor/intellisense>> Visitado por última vez 20.02.2025.
- OpenAI. (s.f.). ChatGPT. <<https://chatgpt.com/>>. Visitado por última vez 20.02.2025.
- Qureshi, B. (2023). Exploring the use of ChatGPT as a tool for learning and assessment in undergraduate computer science curriculum: Opportunities and challenges.
- Sánchez Allende, J. (2023). Aprendiendo a programar con ChatGPT. Edunovatic2023, Conference Proceeding: 8th Virtual International Conference on Education, Innovation and ICT, November 29-30, 2023, pags:369-373.
- Sande, F. de, & López Ramos, P. (2023). El impacto de asistentes basados en IA en la enseñanza-aprendizaje de la programación. Actas de las JENUI 2023.
- Terán, H. (2023). La implementación de la Inteligencia Artificial en la enseñanza de la programación. Un estudio sobre el uso ético de ChatGPT en el aula. Encuentro Internacional de Educación En Ingeniería, 1-11.
- Vaithilingam, P., Zhang, T., & Glassman, E. L. (2022). Expectation vs. Experience: Evaluating the Usability of Code Generation Tools Powered by Large Language Models. CHI Conference On Human Factors In Computing Systems Extended Abstracts