

Impacto de ChatGPT en la enseñanza: Un enfoque de aula invertida para fundamentos de programación

Michael Apaza Calsin¹, Marco Aedo², Eveling Castro³

mapazacals@unsa.edu.pe; maedol@unsa.edu.pe; ecastro@unsa.edu.pe

^{1,2,3} Escuela Profesional de Ingeniería de Sistemas, Universidad Nacional de San Agustín, Santa Catalina 117, C.P. 04001, Arequipa, Perú.

DOI: 10.17013/risti.52.97-112

Resumen: Actualmente las herramientas de Inteligencia Artificial (IA) han surgido de forma disruptiva afectando a campos como el de las tecnologías de la información, específicamente al desarrollo de software, para el cual herramientas como ChatGPT abren nuevas oportunidades y desafíos en la enseñanza y el aprendizaje de conceptos de programación. Actualmente, esta se basa en métodos pedagógicos convencionales y consideramos la adaptación esencial pues estas herramientas de IA cuentan con un potencial que aún está en exploración. Integrar las herramientas de IA en la educación requiere una comprensión profunda de sus capacidades y limitaciones, por eso, en este artículo describimos nuestra experiencia en la adopción de un enfoque de aula invertida para la enseñanza de tópicos básicos de programación aplicada sobre estudiantes de primer año de universidad.

Palabras-clave: Inteligencia Artificial, ChatGPT, Educación, Programación

Impact of ChatGPT on Teaching: A Flipped Classroom Approach for Programming Fundamentals

Abstract: Currently, Artificial Intelligence (AI) tools have emerged disruptively, affecting fields such as information technology, specifically software development. Tools like ChatGPT, for instance, open new opportunities and challenges in teaching and learning programming concepts. Currently, education relies on conventional pedagogical methods, but we consider adaptation essential because these AI tools have untapped potential. Integrating AI tools into education requires a profound understanding of their capabilities and limitations. In this article, we describe our experience in adopting a flipped classroom approach for teaching basic programming topics to first-year university students..

Keywords: Artificial Intelligence, ChatGPT, Education, Programming

1. Introducción

En los recientes años se han dado varias mejoras en el campo tecnológico, especialmente en aquellas herramientas disruptivas basadas en inteligencia artificial, como lo es

ChatGPT, un gran modelo de lenguaje (LLM) desarrollado por OpenAI, un ChatBot soportado bajo el modelo GPT-3.5 (Chat Generative Pre-Trained Transformer, por sus siglas en inglés) enfocado en la conversación (Tian, Lu, On, Tang, Cheung, Klein & Bissyandé, 2023), que ha alcanzado un amplio rango de usuarios desde su lanzamiento en 2022, siendo una de las principales herramientas que más han influido y de las que más interés ha causado en los últimos años desde su apertura al público.

La gran aceptación de esta herramienta y la motivación e interés que supuso para las grandes tecnológicas estimuló la inversión de recursos en la elaboración de sus propios modelos, como Bard de Google o LLaMA de Meta. Este avance provocó una ola de preocupación en las personas sobre el impacto que podrían llegar a tener o cómo es que estas herramientas les afectaría en un futuro (Vorobeva et al., 2022), y según (Baidoo-anu & Owusu, 2023) una sensación dividida entre educadores, adoptando este tipo de herramientas como el futuro o tomando una posición escéptica al indicar que podría apaciguar o eliminar las habilidades analíticas de los estudiantes. A pesar de esto, estudios como el de (Zohair, 2018) señalan que estos estudios no tienden a revelar información fiable, puesto que existen barreras para predecir el futuro (trabajo y empleabilidad), proponiendo optar por un enfoque adaptativo sobre uno preventivo, entre estas barreras destacan ideas cómo, dónde deberían de aplicarse, como se realizaría, atender a las necesidades de un mercado en constante cambio y la realidad social y regional (Morgan et al. , 2019).

Con esta idea se han propuesto varias aplicaciones sobre cómo es que pueden afectar las herramientas actuales en distintas áreas como la investigación, el campo social, la educación, etc. Uno de los primeros sectores afectados fue el sector tecnológico, específicamente el del desarrollo de software. Y por supuesto, que en un principio también se comentó que los desarrolladores podrían llegar a ser reemplazados, según (J. Bommarito, M. Bommarito, D. Katz, & J. Katz 2023) estas herramientas podrían incentivar la mejora de las habilidades y conocimiento de los profesionales de la industria. La aparición de herramientas como GitHub Copilot en el 2021 fue uno de los principales factores de interés en la industria, un modelo de inteligencia artificial basado en Codex de OpenAI que se enfoca en el autocompletado de código fuente, capaz de mejorar la productividad de un desarrollador realizando las mismas actividades hasta un 50% más rápido en comparación con un desarrollo sin el apoyo de esta herramienta (Peng, Kalliamvakou, Cihon, & Demirer 2023) y por supuesto el uso de herramientas alternativas como Tabnine, que pueden llegar a completar alrededor de la mitad del código fuente escrito por el desarrollador y permitir que este se enfoque en la lógica del problema a solucionar en vez de detalles como la sintaxis del código (Reini, 2022).

La aparición de estas herramientas ha permitido un gran avance en la industria y también ha revelado la flexibilidad de los involucrados al aceptar estos cambios, evidencia de esto nos la brinda la Stack Overflow Survey de 2023 (StackOverflow 2023), que indica que las herramientas de IA tienen una adopción de aproximadamente la mitad de desarrolladores, siendo utilizada no solamente para el desarrollo, sino que también para el aprendizaje, además que muy probablemente se siga incrementando la cantidad de personas que las utilicen en los años posteriores.

Es por el gran impacto que pueden llegar a tener estas herramientas en un futuro y especialmente en el futuro de los estudiantes de las distintas áreas de TI que consideramos

necesario incluir el uso de estas herramientas en la enseñanza, integrando modelos de IA como ChatGPT como fuentes de información sobre tópicos relacionados al área, además de aprovechar las distintas ventajas que proveen con respecto a herramientas de enseñanza tradicionales. Sin embargo, también es necesario considerar las posibles desventajas de su uso, puesto que, como se señala en (King, 2023) los estudiantes pueden utilizar ChatGPT para hacer trampas, siendo necesario adaptar métodos de enseñanza interactivos que permitan evidenciar un correcto aprendizaje y adicionalmente herramientas de software de detección de plagio.

Hemos decidido utilizar ChatGPT debido a su nivel de accesibilidad, la capacidad de respuesta en distintos idiomas y la generación de respuestas relevantes en el ámbito de la programación. Además de su efectividad a la hora de generar código explicado y entendible para problemas con una lógica sencilla (Nascimento, Alencar & Cowan, 2023), pudiendo ser una herramienta de interés para estudiantes que se introducen a conceptos de programación y desarrollo de software. A través de este trabajo de investigación, presentaremos el efecto que puede llegar a tener la aplicación de ChatGPT con un enfoque de aula invertida en estudiantes de Ingeniería de Sistemas que cursan asignaturas de programación básica e identificar su efectividad a través de una encuesta y la comparación de la calificación obtenida durante ese periodo de tiempo por su rendimiento académico en comparación con la calificación obtenida por los estudiantes de años previos a partir del 2019.

2. Trabajos Relacionados

El impacto de las herramientas de IA se hace notar en la industria del desarrollo y la educación, un enfoque general de su impacto nos lo brindan (Adiguzel, Kaya & Cansu, 2023) quienes realizan una revisión de la literatura existente sobre las herramientas de IA capaces de generar respuestas similares a las de un ser humano basadas en modelos de lenguaje natural, concluyendo que el uso de estas herramientas en la educación podría significar la individualización de la misma, la retroalimentación y apoyo. Resalta también el gran avance que se da en la aplicación de estas herramientas en distintos entornos y la necesidad de conservar a los docentes como medio de apoyo y soporte pues existe la posibilidad de sesgo en estas herramientas. Por otro lado, (Karina, et al., 2020) resalta la necesidad de competencias digitales en el ámbito educativo, pues resultan fundamentales para adaptar a los estudiantes a la era digital, así como también el fortalecimiento del proceso educativo para adaptar estrategias pedagógicas responsables.

Por lo mismo se han dado investigaciones sobre el rendimiento de estas herramientas y eficiencia en distintos entornos, por ejemplo tenemos a los autores de (J. Bommarito, M. Bommarito, D. Katz, & J. Katz 2023), quienes realizan una evaluación de rendimiento de ChatGPT como un trabajador de conocimiento a través del examen CPA (Certified Public Accountants, por sus siglas en inglés), un examen estandarizado donde también evalúan a otros modelos de IA de OpenAI como: text-davinci-001, text-curie-001, text-babbage-001 y text-ada-001. Donde concluyen que estos modelos son capaces de resolver distintos tipos de ejercicios en distintas áreas de conocimiento, sin embargo, tienden a tener problemas cuando tratan problemas relacionados al cálculo.

Los autores de (Tian, Lu, On, Tang, Cheung, Klein & Bissyandé, 2023) nos presentan un enfoque alternativo para evaluar a ChatGPT, estos optan por realizar la evaluación de la herramienta con un enfoque en las actividades realizadas durante el desarrollo de software, seleccionan tres actividades del área: la generación de código, reparación del mismo y sumarización (resumen explicativo) del código, donde los autores concluyen que la herramienta puede realizar de estas actividades de forma efectiva, sin embargo identifican problemas durante la explicación de código, ya que no puede explicar problemas poco comunes y no logra abstraer la lógica general de los problemas que se le plantean. Además de indicar que utilizar entradas demasiado largas pueden resultar contraproducentes para obtener información de interés.

En lo que respecta a la educación, Halaweh (2023) nos brinda información relevante sobre la aplicación de ChatGPT en un entorno educativo de forma responsable, destacando entre sus principales ventajas: la calidad de escritura o el desarrollo de habilidades de investigación. Además, también resalta otras consideraciones a tener en cuenta, como la privacidad de los datos y el constante uso de la información que se le brinda para mejorar el modelo, la posible generación de dependencia de esta herramienta, consideraciones éticas y morales. Señalando además que pueden surgir nuevas consideraciones a tomar en cuenta en un futuro por lo tanto recomienda adoptar un enfoque proactivo y responsable.

Una de las principales herramientas de apoyo a los desarrolladores basadas en IA es GitHub Copilot. Y consideramos importante dar una revisión de la literatura al mismo. Se ha medido su efectividad desde distintos enfoques, (Asare, Nagappan & Asokan, 2023) evalúan la efectividad de esta herramienta en comparación con los seres humanos a la hora de codificar e introducir vulnerabilidades no previstas, dando como resultado que las sugerencias de esta herramienta no siempre resultan efectivas e incluso llegan a catalogarse como inseguras, a pesar de esto no resultan tan ineficientes como lo pueden llegar a ser los seres humanos. Otra evaluación realizada sobre GitHub Copilot es la de los autores de (Yetistiren, Ozsoy & Tuzun, 2022), quienes también evalúan la capacidad de generación de código en función de validez correctitud y eficiencia, donde determinan que puede brindar una solución correcta el 28.7% de las veces para el conjunto de datos que seleccionaron, además la eficiencia del código generado puede ser similar a uno desarrollado por el ser humano.

Varios estudios también se han enfocado en la aplicación de estas herramientas de IA en la educación como lo realiza el autor de (Qureshi, 2023) quien dividió a estudiantes en 2 grupos para la asignación de tareas de codificación, revelando que mientras más complejo era el problema, más tiempo les tomaba a los estudiantes que hacían uso de ChatGPT resolverlos, además de resolver dichos problemas con ciertas inconsistencias e imprecisiones que posteriormente serían difíciles o muy complejos de mantener o modificar. Algunas consideraciones a tomar en cuenta sobre el uso de ChatGPT en la educación nos la mencionan (Baidoo-anu & Owusu, 2023), y es que a pesar de señalar las distintas ventajas que pueden traer estos modelos de IA, una de sus principales desventajas es la falta de interacción humana, aquella que siempre se ha dado entre un docente o tutor y sus alumnos, además de señalar algunas consideraciones éticas a tomar en cuenta para utilizar estas herramientas.

Entre las metodologías disponibles, seleccionamos la de aula invertida, pues resulta flexible y sencilla de utilizar, siendo que (Contreras et al., 2017) la utiliza para evaluar y posteriormente mejorar materiales de aprendizaje, otro ejemplo nos lo brinda (Gren, 2020) quien utiliza encuestas y entrevistas para recopilar datos de los estudiantes, comparando los resultados de los estudiantes que no participaron en el enfoque de aula invertida concluyendo que resulta efectiva, pero no motivadora para los estudiantes, pues su actitud frente a los temas se conserva. También destaca la importancia de la retroalimentación y la necesidad de una correcta capacitación de los docentes con respecto al uso de esta metodología.

Otro estudio similar es dado por los autores de (Knutas, Herala, Vanhala, & Ikonen, 2016) donde aplica el enfoque de aula invertida en 2 cursos de programación distintos, uno enfocado en la enseñanza de POO, y el otro en aprendizaje sobre aplicaciones web. Durante el desarrollo de su investigación resaltan aspectos importantes a tomar en cuenta, como: una fuente adecuada de información, integración del material de aprendizaje con el del curso, soporte a los alumnos en clase, la importancia de la retroalimentación personalizada y la carga de trabajo que supone para el instructor diseñar las actividades, para finalmente concluir que este enfoque resulta más efectivo que el tradicional.

3. Metodología

El trabajo de investigación realizado sigue un diseño cuasiexperimental, donde estudiantes de primer año de universidad utilizaron ChatGPT como herramienta de apoyo para resolver prácticas elaboradas en base a tópicos específicos de programación relacionados con la asignatura de Fundamentos de Programación 1.

Las prácticas fueron aplicadas solamente a un grupo determinado de estudiantes, puesto que el dictado del curso se da entre varios docentes. Para los resultados de esta investigación, solamente se considera a los estudiantes que fueron asignados al docente que permitió el desarrollo de esta actividad. Para realizar la comparación de notas, se utilizaron las calificaciones académicas obtenidas por estos estudiantes con las de estudiantes de años previos que también trataron los mismos tópicos y con el mismo docente.

3.1. Población y muestra

Para realizar el estudio se tomó en cuenta a 26 estudiantes de primer año de Ingeniería de Sistemas a los que se les asignaron prácticas de refuerzo y se les indicó que debían de utilizar ChatGPT para solucionar los problemas de programación y además de completar un cuestionario sobre la actividad. Los tópicos seleccionados para la elaboración de las prácticas fueron los seleccionados en base al avance académico de las últimas etapas del curso, estos fueron: arreglos unidimensionales, arreglos bidimensionales e introducción a la Programación Orientada a Objetos (POO). Para el desarrollo de las actividades se optó por adoptar un enfoque de aula invertida (véase la figura 1). Esto implicó presentar a los estudiantes conceptos que no habían sido vistos en clases hasta el momento. Esto brindó a los estudiantes la oportunidad de interactuar con nuevos conceptos por adelantado, lo que podría mejorar su comprensión cuando se aborden durante la clase tradicional.

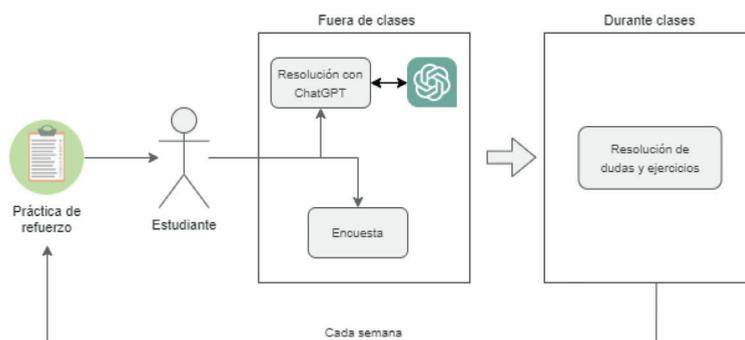


Figura 1 – Enfoque de aula invertida utilizando ChatGPT

También, debido a que las prácticas podrían ser resueltas simplemente utilizando la herramienta, la necesidad de garantizar la obtención de información relevante, la retención y entendimiento de los tópicos tratados, se solicitó a los estudiantes que utilicen los términos y conceptos aprendidos en el curso hasta el momento para explicar la solución obtenida. Asegurando que los estudiantes utilicen no solamente la herramienta para dar resolución a los problemas, sino que utilicen el conocimiento adquirido en los tópicos previos.

3.2. Procedimiento

A lo largo de tres semanas se prepararon prácticas que buscaban introducir a 26 estudiantes al siguiente tema según el cronograma académico planificado para el curso, el objetivo fue buscar un enfoque de aula invertida donde los estudiantes obtendrían información previa a las clases dictadas utilizando ChatGPT, para esto, se les asignaron prácticas durante las tres semanas de actividades académicas, se motivó a elaborar las consultas sobre los temas que desconocían y con el objetivo de resolver los ejercicios planteados, para al final solicitar completar una encuesta para conocer la opinión del estudiante sobre el método, si es que considera que logró la meta de aprendizaje e identificar la cantidad de consultas que el estudiante realizó en promedio por cada ejercicio. Las respuestas obtenidas para cada encuesta fueron reduciendo conforme se desarrollaba la investigación, se recibieron un total de 26 respuestas de los estudiantes en la primera sesión, en la segunda se recibieron 16 respuestas, para finalmente registrar un total de 11 respuestas en la última encuesta.

Para evaluar el nivel de mejora de los estudiantes, se realizó la comparación de las calificaciones obtenidas por los estudiantes en comparación con las obtenidas por estudiantes de años previos (2019-2022), tomando en cuenta los dos tipos de evaluación, además de utilizar la escala de 5 valores que se utilizan en los informes de desempeño en la institución educativa. Las 2 formas de evaluación que se dan en la institución son: evaluación continua y evaluación por examen, en un semestre académico normal se realizan un total de 6 evaluaciones y para esta investigación solo se consideran las 2 últimas de cada año, pues estas son las que representan las notas asignadas a los tópicos seleccionados para la investigación.

- Evaluación continua (EC): Evaluación realizada por cada etapa durante todo el semestre académico, tales como actividades, prácticas de laboratorio, evaluaciones en clase, etc.
- Evaluación parcial (EX): Examen de conocimientos realizado por cada etapa sobre los temas realizados.

3.3. Calificaciones

Como se mencionó anteriormente en esta sección se analizarán los resultados de las calificaciones de los 4 años previos en el curso de Fundamentos de Programación 1, considerando solamente las calificaciones relacionadas a los tópicos seleccionados para este trabajo, tomando calificaciones desde la enseñanza presencial en 2019 hasta el presente año. La institución toma un rango de valores para la calificación desde 0 (cero) para la nota mínima y 20 como máxima. Siendo las notas a entregar las antes mencionadas EC y EX.

En la tabla 1, se muestran las calificaciones de los estudiantes durante su tercera etapa entre los años 2019 y 2023, consideramos necesario utilizar un valor porcentual debido al cambio en la cantidad de estudiantes a lo largo de los años. Se representan ambos tipos de evaluación y el total de estudiantes por año. Los datos obtenidos para el año 2023 son los que representan los resultados de la aplicación de ChatGPT y un enfoque de aula invertida en los estudiantes, En la siguiente sección se explicarán los resultados obtenidos durante esta investigación, considerando principalmente los resultados de la evaluación EX, puesto que son relevantes para evaluar el correcto aprendizaje por parte de los estudiantes, sin embargo también se comentará sobre la calificación obtenida en la EC y cómo varía con respecto a años previos.

| Rango nota | 2019 | | 2020 | | 2021 | | 2022 | | 2023 | | Calificativo |
|---------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|-----------------|
| | EC | EX | |
| AB | 14.75% | 15.57% | 5.95% | 13.10% | 13.64% | 13.64% | 13.01% | 16.26% | 6.98% | 6.98% | Abandono |
| 1-7 | 2.46% | 37.70% | 9.52% | 17.86% | 6.36% | 30.91% | 7.32% | 14.63% | 18.60% | 23.26% | Insatisfactorio |
| 8-10 | 6.56% | 18.85% | 4.76% | 16.67% | 3.64% | 17.27% | 7.32% | 15.45% | 6.98% | 23.26% | En proceso |
| 11-14 | 27.87% | 18.03% | 35.71% | 29.76% | 26.36% | 22.73% | 25.20% | 35.77% | 30.23% | 25.58% | Satisfactorio |
| 15-20 | 48.36% | 9.84% | 44.05% | 22.62% | 50.00% | 15.45% | 47.15% | 17.89% | 37.21% | 20.93% | Sobresaliente |

Tabla 1 – Distribución porcentual de estudiantes en el 2023

Para brindar un contexto aproximado sobre la media de calificaciones obtenidas en la institución, se considera presentar las notas mínimas, máximas y el promedio de las notas de todos los estudiantes entre los años 2019 y 2023, observe la figura 2 y la figura 3 para observar las notas de EC y EX respectivamente entre los años mencionados.

El mayor promedio de calificación observado en este estudio para la EC se obtuvo en el 2019 siendo de 14.49, que posteriormente se fue reduciendo mínimamente hasta la fecha. Mientras que la menor media de EX se obtiene también en este mismo año e incrementa durante los siguientes años. Este cambio a través del tiempo puede deberse

a varios factores, desde el nivel de rigurosidad de la calificación, el docente asignado al curso, la transición a la enseñanza remota, etc. Para esta investigación reconocemos la existencia de estos factores, sin embargo, nos enfocamos en las calificaciones de los estudiantes disponibles y no en los factores que pudieron modificarlos.

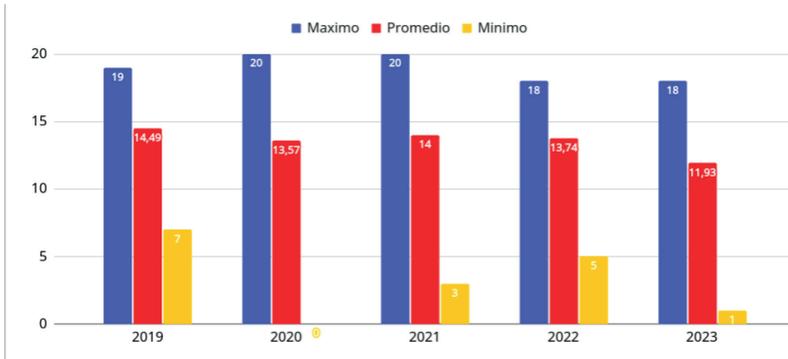


Figura 2 – Calificaciones de evaluación continua 2019 - 2023

En el caso de la media de evaluación continua, resaltamos una tendencia en el cambio de notas desde el año 2019, puesto que se da un decremento medianamente notorio en las calificaciones de los estudiantes (normalmente superando los 13.5 en años anteriores y obteniendo 11.93 en el 2023), esto puede deberse a la transición existente entre la enseñanza presencial y la remota, puesto que este cambio en la media también es notorio en la evaluación realizada mediante exámenes. Otro factor de importancia puede ser el cambio en el porcentaje de la nota final (calificación consolidada de las 6 notas obtenidas a lo largo del semestre), ya que para asegurar un correcto aprendizaje y desempeño, se realizó un cambio en la proporción de las notas, pudiendo suponer un mayor esfuerzo por obtener mejores notas en la EC que en la EX por parte de los estudiantes.

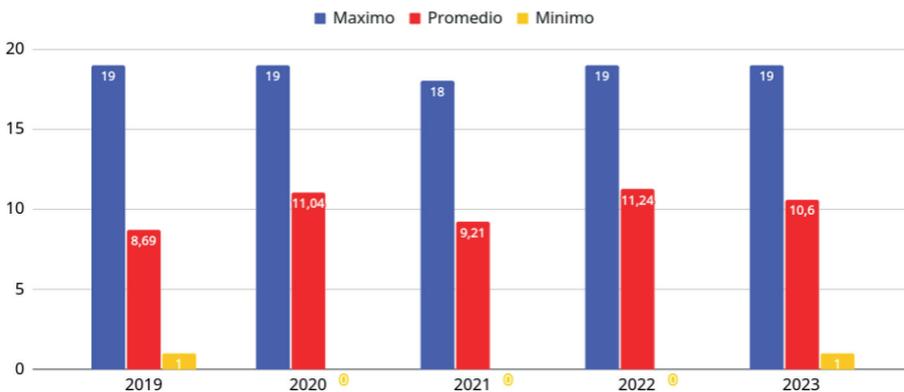


Figura 3 – Calificaciones obtenidas en los exámenes 2019 - 2023

Como se indicó anteriormente, en el caso de la media de la calificación mediante la evaluación por examen se distingue un cambio notorio, pues esta se reduce (siendo 14.49 en el 2019 y 11.04 en el 2020), y se conserva el cambio durante los años de pandemia tanto en el caso de la evaluación continua como por examen.

3.4. Resultados de las encuestas

Junto a cada práctica desarrollada, a lo largo de cada semana se han brindado encuestas a los estudiantes para obtener su opinión sobre su experiencia realizando esta actividad, a través de las siguientes preguntas:

- ¿Qué opinión tiene sobre el método de enseñanza aplicado en la actividad de avance? Calificado desde 1 (excelente) a 5 (deficiente)
- ¿Cuántas consultas por pregunta realizó a ChatGPT en promedio? Calificado con: 1, 2 o 3, 4 o 5 y más de 5.
- ¿Tiene alguna observación? Opinión no obligatoria del estudiante.

En la figura 4, se pueden observar los resultados de la encuesta, el interés de los estudiantes y su disposición para optar por un nuevo método de enseñanza se hace evidente, ya que ninguno consideró malo o pésimo el método de enseñanza, sin embargo, son varios los que consideran la forma de enseñanza algo regular, esto puede deberse a factores como la eficiencia de la herramienta, la sobreactividad que supone la adición de estos ejercicios a los estudiantes o simplemente la facilidad que brinda tener un docente en una clase presencial con quien atender las dudas que pudiera tener, explicaremos a mayor profundidad este punto al final de esta sección en base a las observaciones obtenidas por las mismas encuestas.

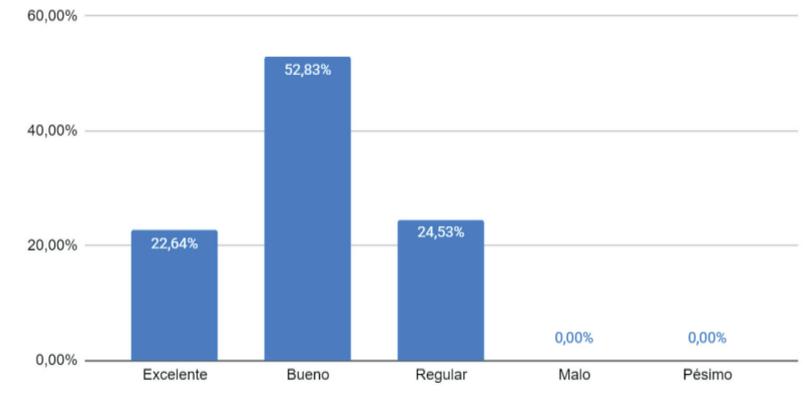


Figura 4 – Opinión de los estudiantes sobre la metodología aplicada.

El segundo punto de interés fue el saber si es que la herramienta resultó eficiente al punto de lograr enseñar sobre el tópico al estudiante, este se midió a través de 3 alternativas: si logro entender, más o menos, o no entendió. para lo cual se obtuvieron los resultados mostrados en la figura 5.

Más de la mitad de los estudiantes consideran que lograron comprender el tópico presentado, mientras que cierta cantidad no considera que lo haya entendido al completo y a su vez ninguno considera que no haya entendido nada, esto puede deberse a que para cada práctica se ha brindado sugerencias sobre qué debería de solicitar para obtener una respuesta entendible, y es que como mínimo se espera que se entienda un concepto básico del tema a tratar.

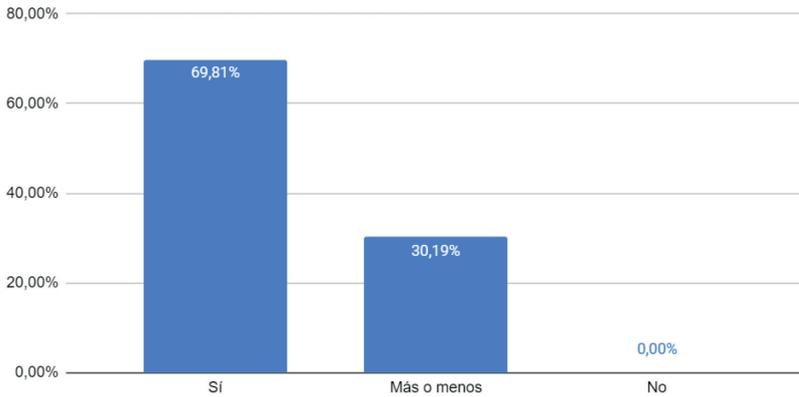


Figura 5 – Calificación sobre el entendimiento del tema.

Finalmente, la tercera pregunta estaba dirigida a identificar la cantidad de consultas que los estudiantes realizan por cada ejercicio, pues puede resultar relevante conocer si es que los estudiantes son capaces de elaborar instrucciones correctamente estructuradas o con la especificidad correcta para obtener respuestas satisfactorias a las consultas que realizaron, los resultados se pueden observar en la figura 6.

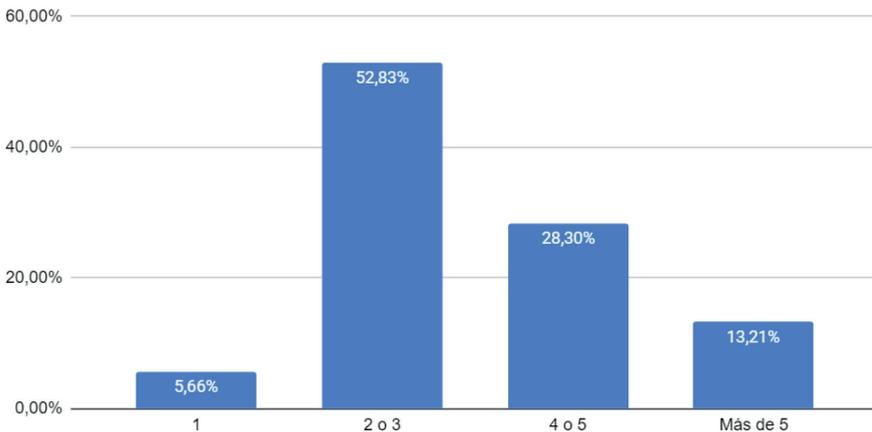


Figura 6 – Cantidad de consultas realizadas por cada ejercicio.

Esta estadística revela que la gran mayoría de estudiantes resuelve sus dudas con 2 o 3 consultas, seguido por aquellos que las resuelven en 4 o 5, para finalmente ser una minoría aquellos que las resuelven con 1 o más de 5 consultas. Estos resultados pueden deberse a la falta de experiencia con el uso de este tipo de herramientas, pues estas no funcionan de la misma manera que los buscadores en internet, y requieren un mayor contexto en la formulación de las entradas para obtener una salida útil para el usuario (White, Fu, Hays, Sandborn, Olea, Gilbert, Elnashar, Spencer-Smith & Schmidt, 2023).

3.5. Sobre los datos

El grado de mejora académica se dará en base a los datos de las calificaciones anteriores de estudiantes de años previos que cursaron la misma etapa de evaluación, considerando los 4 años previos a este estudio, además de esto, cabe recalcar que las notas obtenidas desde el año 2020 pudieron ser afectadas debido a que la institución adoptó un método de enseñanza remota, sin embargo, también se consideraron las calificaciones obtenidas durante el año 2019, año en el que se realizaron las actividades académicas con normalidad.

Además de la situación de enseñanza, también se encontraron desafíos en la obtención de la información previa al año 2019, puesto que el método de calificación planteada por la institución en años previos no se adecúa a los datos obtenidos a partir del año mencionado hasta la actualidad, dificultando la obtención de notas de los tópicos seleccionados para este trabajo de investigación, razón por la cual no se contemplaron estas calificaciones en este estudio.

4. Resultados

En esta sección brindaremos los resultados de la investigación y la metodología aplicada a los estudiantes de Ingeniería de Sistemas, además de comparar su rendimiento académico con respecto a años previos, además de resumir los resultados de las encuestas.

La media de los resultados de los estudiantes a través de los años indica que se tenía cierta tendencia a mantenerse hasta la fecha de la presente investigación, el cambio más relevante se da durante los años trabajados de manera remota, aunque puede resultar relevante, en cierta medida se mantiene equilibrada (puesto que se dio un incremento en la media de la nota continua de los estudiantes y un decremento en la media de la calificación por examen). No se ha identificado una mejora considerable entre la media de las calificaciones obtenidas en el año 2023 en comparación con el resto de años, sin embargo, no descartamos la presencia de datos de interés.

La distribución de los estudiantes en base a su calificación por EX (véase la figura 7) a lo largo de los años ha sido variada, siendo generalmente las calificaciones “satisfactorias” e “insatisfactorias” las que ocupaban los picos más altos (mayores al 30%), sin embargo, en este último año cada categoría de calificación, a excepción de los estudiantes en situación de abandono, representan entre el 20.93% y 25.58% del total de estudiantes. En la tabla 2 representamos la diferencia en la proporción de estudiantes por cada

calificación, representamos el cambio de este último año en comparación con la media de los años previos.

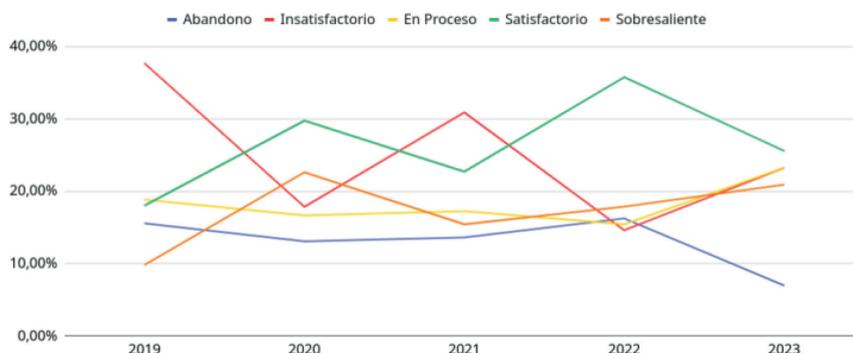


Figura 7 – Proporción de estudiantes según calificaciones de EX entre el 2019 y 2023

| Calificación | Promedio (2019-2022) | 2023 | Diferencia |
|-----------------|----------------------|--------|------------|
| Abandono | 14.64% | 6.97% | -7.66% |
| Insatisfactorio | 25.28% | 23.26% | -2.02% |
| En proceso | 17.06% | 23.26% | +6.20% |
| Satisfactorio | 26.57% | 25.58% | -0.99% |
| Sobresaliente | 16.45% | 20.93% | +4.48% |

Tabla 2 – Diferencia porcentual de calificaciones EX entre la media de 2019-2022 y 2023

Durante la investigación se aplicó una encuesta a los estudiantes sobre el uso de esta herramienta como fuente de información, la metodología y su experiencia durante su aplicación, a continuación, se enumeran los problemas más frecuentes que los estudiantes identificaron:

- Las consultas realizadas a ChatGPT pueden resultar complejas puesto que no siempre obtienen resultados que ayuden a entender los conceptos propuestos, sino que pueden resultar en respuestas complejas, incluir conceptos avanzados o simplemente brindar información que no resulta relevante.
- Las respuestas no resultan claras para el estudiante, ya que indicaron que fue necesario realizar varias consultas o elaborar una bastante específica para obtener información relevante.
- La metodología aplicada no resultó eficiente debido al tiempo de resolución que se les brindó para resolver la práctica y el desconocimiento del tema sobre el cual trata dicha práctica.

Las opiniones de los estudiantes con respecto al método de enseñanza revela una efectividad media, puesto que si bien la mayoría indican que es “Bueno” puede que los problemas mencionados anteriormente sean los causantes de no obtener una calificación

“Excelente”, también, el total de estudiantes que respondieron la encuesta considera que ha logrado aprender al menos algo sobre el tema, finalmente la mayor frecuencia en la cantidad de consultas que realizaron a ChatGPT por cada ejercicio está entre 2 o 3.

5. Discusión

Las calificaciones obtenidas no revelaron cambios relevantes en la tendencia de las puntuaciones estudiantiles a lo largo de los años. Sin embargo, se identificó la posibilidad de mejorar el desempeño de aquellos estudiantes que muestran dificultades en su rendimiento académico. El uso de ChatGPT como fuente de conocimiento acompañado de un enfoque de aula invertida ha demostrado tener un impacto considerable entre los alumnos menos destacados, sin causar una perturbación sustancial en el rendimiento del resto de los estudiantes. Además, es esencial considerar una ampliación del alcance de este estudio. Esto implicaría analizar una extensa serie de datos históricos recopilados en un contexto similar para identificar tendencias más arraigadas.

Es importante mencionar que el uso de calificaciones obtenidas durante la pandemia puede tener un impacto considerable en los resultados de la investigación y, por tanto, es necesario abordar este aspecto con mayor profundidad en futuros análisis. Además, para trabajos futuros se recomienda realizar un estudio más duradero y dar un mayor seguimiento a los resultados de cada estudiante, puesto que debido a la planificación de la institución solamente se evaluó a los estudiantes durante parte del ciclo académico y se limitó el alcance de la evaluación a las calificaciones académicas.

Por otro lado, los resultados de las encuestas indican que los estudiantes consideran útil el método de enseñanza, e indican problemas como: la cantidad de actividades a realizar, la especificidad requerida para obtener resultados relevantes e incluso la percepción de dificultad en el aprendizaje o la claridad de las respuestas obtenidas de ChatGPT. Cabe recalcar que, si bien no se dio en esta encuesta, también es necesario considerar que las recomendaciones de los estudiantes pueden dejar de lado tópicos importantes que se tienen que tocar en el curso (Knutas, Herala, Vanhala & Ikonen, 2016).

Fuera del foco de la investigación, también se identifica cierta variación en las notas entre los años 2019 y 2020 ya que la evaluación a los estudiantes durante los años afectados por la pandemia y la modalidad de enseñanza remota afecta mínimamente su rendimiento, siendo reducida la media cuando se trata de EC e incrementando cuando se trata de EX. Con respecto a la evaluación continua, consideramos que se denota un crecimiento debido a la facilidad de investigación y trabajo desde el hogar, sin embargo, esta comodidad también afecta al rendimiento y la adquisición de conocimientos siendo evidencia de esto la media de notas del examen durante estos años.

6. Conclusiones

En este estudio se motivó a estudiantes de Ingeniería de Sistemas a que hicieran uso de ChatGPT para la resolución de ejercicios de programación con una base previa sobre conceptos básicos de programación tales como: variables, librería Math de Java, bucles y condicionales, etc. Las actividades se realizaron con un enfoque de aula invertida, brindándoles prácticas de programación en base a los tópicos académicos asignados a

su etapa, además de una encuesta para obtener información sobre su experiencia con el desarrollo de la actividad y su opinión sobre el método utilizado. Para finalmente analizar los resultados obtenidos en base a su calificación académica en comparación con los obtenidos por estudiantes de años previos.

Las calificaciones obtenidas para el periodo seleccionado (2023) no indican una mejora relevante en todos los estudiantes, sin embargo, si se sale a relucir una tendencia en las notas obtenidas de los años previos con respecto a la evaluación de conocimiento (EX), se puede identificar una mejora en la distribución de los estudiantes con respecto a sus calificaciones. Viéndose reducida considerablemente la cantidad de estudiantes que tienden a obtener notas bajas con respecto a la media de los años anteriores.

La opinión de los estudiantes sobre la actividad fue positiva en su mayoría, sin embargo, considerando las observaciones obtenidas, se recomienda adoptar algunas medidas, como una capacitación previa sobre el enfoque de aula invertida, la reducción de la cantidad de actividades sin comprometer el objetivo de aprendizaje e incluso también considerar una capacitación previa a los estudiantes sobre la estructuración de consultas a los grandes modelos de lenguaje como ChatGPT. Además, los estudiantes consideran que han logrado los objetivos de aprendizaje de las prácticas ya que, en las respuestas registradas, la mayoría indica que comprendió el tópico asignado.

En relación al uso de ChatGPT como fuente de información para conceptos de programación, la consideramos una herramienta idónea por su nivel de accesibilidad y su capacidad de respuesta, pues es capaz entender cuestiones de distintas complejidades y ofrecer la información pertinente respecto al tema. No obstante, es importante señalar que la estructura de las consultas requiere un conocimiento previo de la herramienta. Como ya mencionamos, esta limitación puede ser superada mediante una capacitación preliminar de los usuarios o a través de la exploración de otros modelos de lenguaje o implementaciones específicas centradas en la resolución de consultas relacionadas exclusivamente con la programación.

7. Trabajos futuros

Para finalizar, en base a estos resultados obtenidos, como trabajo futuro se plantea incluir una capacitación sobre la adopción de técnicas de “Prompt Engineering” para mejorar la búsqueda de resultados en modelos de lenguaje considerando aspectos como las limitaciones, contextos, foco de la consulta, etc. Esto debido a que la opinión de los estudiantes revela que consideran útil esta forma de enseñanza, pero requieren mayor experiencia con el uso de esta herramienta ya que muchas veces las consultas realizadas por los estudiantes no resultan útiles debido a entradas confusas, respuestas demasiado avanzadas o el hecho de realizar varias consultas con especificaciones cada vez más exactas para obtener una respuesta útil. Además, contemplamos estudiar la aplicación de otros modelos de lenguaje liberados hasta la fecha, con distintos enfoques, permitiendo así evaluar en base a qué combinación de estos se pueden dar una solución óptima a las necesidades de los estudiantes en lo que respecta a su aprendizaje enfocado en conceptos de programación.

Referencias

- Adiguzel, T., Kaya, H., & Cansu, F. (2023). Revolutionizing education with AI: Exploring the transformative potential of ChatGPT. *Contemporary Educational Technology*, 15, ep429. <https://doi.org/10.30935/cedtech/13152>
- Asare, O., Nagappan, M., & Asokan, N. (2023). Is GitHub's Copilot as bad as humans at introducing vulnerabilities in code?. *Empirical Software Engineering*, 28. <https://doi.org/10.48550/arXiv.2204.04741>
- Baidoo-anu, D., & Owusu Ansah, L. (2023). Education in the Era of Generative Artificial Intelligence (AI): Understanding the Potential Benefits of ChatGPT in Promoting Teaching and Learning. *Journal of AI*, 7(1), 52-62. <https://dx.doi.org/10.2139/ssrn.4337484>
- Basit Q. (2023). Exploring the Use of ChatGPT as a Tool for Learning and Assessment in Undergraduate Computer Science Curriculum: Opportunities and Challenges. <https://doi.org/10.48550/arXiv.2304.11214>
- Bommarito, J., Bommarito, M., Katz, D. M., & Katz, J. (2023). GPT as Knowledge Worker: A Zero-Shot Evaluation of (AI)CPA Capabilities. Disponible en: <https://doi.org/10.48550/arXiv.2301.04408>
- Contreras, J. Á., Masa, J. A., Andrade, M. G. M., & Espada, R. M.. (2017). Uso del modelo de aprendizaje inverso para mejorar materiales educativos universitarios. *RISTI - Revista Ibérica de Sistemas e Tecnologias de Informação*. 17-32. [10.17013/risti.23.17-32](https://doi.org/10.17013/risti.23.17-32)
- Daysi, F., & Bedon, A. (2020). Competencias Digitales y Educación: Aproximaciones a los Consumos de Estudiantes en Formación Docente. *RISTI - Revista Ibérica de Sistemas e Tecnologias de Informação*, E31, 318-327.
- Gren, L. (2020). A Flipped Classroom Approach to Teaching Empirical Software Engineering. *IEEE Transactions on Education*, 63(3), 155-163. <https://doi.org/10.1109/TE.2019.2960264>
- Halaweh, M. (2023). ChatGPT in education: Strategies for responsible implementation. *Contemporary Educational Technology*, 15. <https://doi.org/10.30935/cedtech/13036>
- King, M. (2023). A Conversation on Artificial Intelligence, Chatbots, and Plagiarism in Higher Education. *Cellular and Molecular Bioengineering*, 16, 1-2. <https://doi.org/10.1007/s12195-022-00754-8>
- Knutas, A., Herala, A., Vanhala, E., & Ikonen, J. (2016). The Flipped Classroom Method: Lessons Learned from Flipping Two Programming Courses. In *Proceedings of the 17th International Conference on Computer Systems and Technologies 2016* (pp. 423-430). Association for Computing Machinery. <https://doi.org/10.1145/2983468.2983524>

- Mahmoud Abu Zohair, L. (2018). The Future of Software Engineering by 2050s: Will AI Replace Software Engineers?. *International Journal of Information Technology and Language Studies*, 2, 1-13. Disponible en: <https://journals.sfu.ca/ijitls/index.php/ijitls/article/view/23>.
- Morgan, R., Frank, D., Bessen, J. E., Brynjolfsson, E., Cebrian, M., Deming, D. J., Feldman, M., Groh, M., Lobo, J., Moro, E., Wang, D., Youn, H., & Rahwan, I. (2019). Toward understanding the impact of artificial intelligence on labor. *Proceedings of the National Academy of Sciences*, 116(14), 6531-6539. <https://doi.org/10.1073/pnas.1900949116>
- Nascimento, N., Alencar, P., & Cowan, D. (2023). Comparing Software Developers with ChatGPT: An Empirical Investigation. <https://doi.org/10.48550/arXiv.2305.11837>
- Peng, S., Kalliamvakou, E., Cihon, P., & Demirer, M. (2023). The Impact of AI on Developer Productivity: Evidence from GitHub Copilot. <https://doi.org/10.48550/arXiv.2302.06590>
- Reini, N. (2022). The Impact of AI powered code completion in the software engineering field. (Tesis de bachillerato en Administración de Empresas, Business Information Technology, Haaga-Helia University of Applied Sciences). Theseus. Disponible en: <https://www.theseus.fi/handle/10024/781700>.
- StackOverflow. (2023). 2023 Developer Survey. Disponible en: <https://survey.stackoverflow.co/2023/>
- Tian, H., Lu, W., Li, T. O., Tang, X., Cheung, S. C., Klein, J., & Bissyandé, T. F. (2023). Is ChatGPT the Ultimate Programming Assistant – How far is it?. <https://doi.org/10.48550/arXiv.2304.11938>
- Vorobeva, D., Fassi, Y., Costa Pinto, D., Hildebrand, D., Herter, M., & Mattila, A. (2022). Thinking Skills Don't Protect Service Workers from Replacement by Artificial Intelligence. *Journal of Service Research*, 25, 109467052211043. <https://doi.org/10.1177/10946705221104312>
- Yetistiren, B., Ozsoy, I., & Tuzun, E. (2022). Assessing the Quality of GitHub Copilot's Code Generation. In *Proceedings of the 18th International Conference on Predictive Models and Data Analytics in Software Engineering* (pp. 62–71). Association for Computing Machinery. <https://doi.org/10.1145/3558489.3559072>
- White, J., Fu, Q., Hays, S., Sandborn, M., Olea, C., Gilbert, H., Elnashar, A., Spencer-Smith, J., & Schmidt, D. C. (2023). A Prompt Pattern Catalog to Enhance Prompt Engineering with ChatGPT. <https://doi.org/10.48550/arXiv.2302.11382>