

Experiencia docente preliminar con ChatGPT: desafíos y adaptaciones

Francisco P. Romero, Jesus Serrano-Guerrero, Julio Alberto López-Gómez,
Luis Jiménez Linares, José Ángel Martín-Baos

Escuela Superior de Informática. Escuela de Ingeniería Minera e Industrial de Almadén
Universidad de Castilla la Mancha

FranciscoP.Romero@uclm.es Jesus.Serrano@uclm.es
JulioAlberto.Lopez@uclm.es, Luis.Jimenez@uclm.es
joseangel.martin@uclm.es

Resumen

La aparición y amplia difusión de ChatGPT está teniendo un impacto significativo en todos los aspectos de la sociedad, especialmente en el ámbito educativo, incluyendo la educación universitaria. La utilización de herramientas de generación de lenguaje altamente potentes plantea un desafío en la adaptación de las metodologías y actividades dentro del proceso de enseñanza-aprendizaje, ya que suponen un cambio de paradigma en tareas como la codificación o la elaboración de resúmenes. Se han llevado a cabo una experiencia docente en las que se ha adaptado la actividad planificada a la nueva realidad de trabajar con ChatGPT, modificando los objetivos y métodos de evaluación, siempre dentro del marco de los resultados de aprendizaje establecidos. Esta experiencia ha permitido extraer diferentes conclusiones que podrían ser útiles para establecer los siguientes pasos a seguir dentro de un nuevo contexto educativo, considerando problemas como la disponibilidad y la precisión de las respuestas obtenidas.

Abstract

The widespread use of ChatGPT is significantly impacting society, particularly in education. Language generation tools like ChatGPT pose a challenge in adapting traditional teaching methodologies to this new technology, as they represent a paradigm shift in tasks such as coding and summarization. An experiment has been conducted to adapt teaching activities to ChatGPT, modifying objectives and assessment methods while adhering to established learning outcomes. The experiment yielded practical conclusions for selecting the next steps in this new educational context while addressing issues such as the availability and accuracy of ChatGPT-generated answers.

Palabras clave

ChatGPT, Generadores de lenguaje, Metodologías.

1. Contexto

Los modelos generadores del lenguaje pre-entrenados, los denominados GPT [1], y una de sus últimas versiones ChatGPT, han revolucionado muchos aspectos de la vida docente dada su capacidad de entender cuestiones y generar textos a partir de simples preguntas o especificaciones. Los actuales sistemas anti-plagio no son capaces de detectar este lenguaje no elaborado de forma original dado que es lenguaje generado, es decir, el modelo crea respuestas individuales cada vez, no ofreciendo respuestas idénticas en ningún momento. Esta novedosa situación plantea verdaderos interrogantes sobre la forma en que se puede certificar si un estudiante realiza alguna de las tareas de clase.

Dada esta situación, las instituciones educativas pueden orientarse hacia ignorar la existencia de estas herramientas o su prohibición. Esto supone conservar los enunciados y formulaciones de las tareas incorporando un mensaje de prohibición del uso de ChatGPT para su realización. Las condiciones de realización también podrían cambiar dado que algunas tareas autónomas tendrían que pasar a ser presenciales [6]. También supondría incorporar herramientas que detectan el uso de modelos generadores del lenguaje, las cuales aún no tienen la precisión suficiente para su aplicación en una actividad real. De esta forma se plantea la incorporación de estas herramientas a los procesos de enseñanza y aprendizaje, lo cual, puede suponer la aparición de nuevos métodos para evaluar capacidades como el pensamiento crítico, en análisis de la información y la formulación estructurada de preguntas [2].

En este trabajo se presenta un caso de estudio llevado a cabo en la asignatura de *Técnicas de Aprendizaje Automático*, posteriormente se presentan las adaptaciones necesarias que se consideran relevantes para introducir este tipo de herramientas en la docencia de diferentes asignaturas así como los desafíos que presentarán estas adaptaciones. Se finalizará con las conclusiones de toda la experiencia docente.

2. Actividad

Dentro de las actividades previstas en la asignatura '*Técnicas de Aprendizaje Automático*' existe un ítem cuyo nombre es '*Elaboración de informes o trabajos*'. Este tipo de actividades consisten en la elaboración, individual o por parejas, de un '*informe sobre un tema expuesto por el profesor*'. Los enunciados de estas tareas, así como su cariz individual o grupal quedan definidos a principios de Septiembre para su realización durante el primer cuatrimestre. En el momento en el que a finales de Noviembre OpenAI lanza ChatGPT, la tarea que estaba publicada dentro de este grupo consistía en el análisis teórico de diferentes temas relacionados con las tendencias más actuales en la materia, con un enunciado como el que sigue:

Elaborar un texto resumen de máximo 1 página sobre alguno de los siguientes temas tendencia en el campo del aprendizaje automático: TinyML, MLOps, AutoML, Conformal Prediction, Causality

Los criterios de evaluación estaban basados en su contenido, su originalidad y la cantidad y relevancia de las referencias utilizadas. Este enunciado puede ser resuelto fácilmente con una única consulta a ChatGPT, salvo el aspecto de las referencias. Esta situación obliga a hacer una recomposición de la tarea manteniendo el objetivo original de la misma, adquirir conocimientos sobre temas que sean tendencias actuales en el dominio de trabajo, pero introduciendo a la propia herramienta ChatGPT para su elaboración. El enunciado modificado se muestra a continuación.

Mantener una conversación con ChatGPT en el cual se permita profundizar en alguno de los siguientes conceptos [...los mismos que el enunciado antiguo..]. La conversación debe contener al menos cinco preguntas y debe cerrar con unas conclusiones personales sobre los resultados y la conversación misma".

2.1. Evaluación

La evaluación de la tarea se realizó mediante una lista que comprendía diez ítems que permiten evaluar la

conversación desde diferentes puntos de vista. En primer lugar, se consideran aspectos de nivel básico como el foco de la conversación y el número de cuestiones utilizadas. En segundo lugar se evalúa la calidad de las preguntas planteadas que deben de ser específicas, concisas y fáciles de entender, evitando redundancias y ambigüedades, lo cual nos debe permitir extraer la máxima información al modelo interrogado. Finalmente, y como temas avanzados, se busca que se haya conseguido profundizar en el tema, e incluso utilizar las propias respuestas de ChatGPT para elaborar unas mejores preguntas que continúen con la conversación. A su vez, se revisa si se ha verificado que las respuestas dadas, y las referencias, si las ofreciera, fueran veraces.

1. Conversación centrada en el tema específico.
2. Conversación suficientemente larga
3. Las preguntas son específicas y no redundantes.
4. Las preguntas son concisas y directas.
5. Las instrucciones no ambiguas . son fáciles de entender, sin ambigüedades.
6. Las instrucciones son específicas y expresan claramente el tema.
7. La conversación profundiza en los detalles.
8. Al menos dos *prompts* extraídos de las respuestas.
9. El informe contiene conclusiones personales sobre los resultados obtenidos.
10. Si la conversación ofrece referencias se ha comprobado si son veraces.

2.2. Resultados

La tarea estuvo disponible para su realización durante una quincena, pero fue resuelta por los estudiantes en los primeros tres días de ser publicada. Algo excepcional pero que está dentro de la lógica dado que el atractivo de la herramienta resultó ser un gran elemento motivador para la realización del trabajo. El análisis de las respuestas entregadas nos permitió identificar de forma sencilla cuatro tipos de comportamientos ante el uso de estas herramientas que van desde el más básico, compuesto por preguntas o peticiones genéricas, hasta un nivel de excelencia en el cual se aprovecha al máximo la información proporcionada por ChatGPT para profundizar sobre el concepto en cuestión.

- Principiante: Se pregunta sobre un concepto y luego lo único que hacemos es solicitarle más información.
- Básico: Se sigue un guión genérico y parcialmente estructurado, por ejemplo: definición del concepto, ejemplos, ventajas/desventajas, etc.
- Intermedio: Se conoce o se ha preparado el tema, se conocen cuestiones de más difícil solución y con ello preparamos preguntas que cubran esas cuestiones. se sigue un guión más complejo y específico que el anterior.

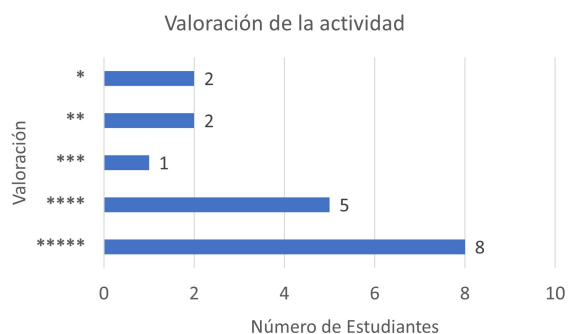


Figura 1: Resultados Encuesta Valoración

- **Avanzado:** Además de lo especificado en el nivel anterior, se aprovechan las respuestas del sistema para profundizar más en el tema, modificando, adaptando y profundizando en el guión ayudándose de las respuestas del propio sistema.

Al finalizar la tarea se planteó recopilar la opinión de los estudiantes, por lo que se lanzó una encuesta para valorar la actividad de manera voluntaria en una escala de 1 a 5. Se recogieron 17 encuestas y la valoración resultó ser muy positiva.

3. Adaptaciones

Como lección aprendida de la actividad planteada se plantean diferentes cambios en la docencia de las asignaturas que van desde el punto de vista competencial así como las diferentes estrategias para mejorar las capacidades de interrogatorio de nuestros estudiantes a partir del uso de este tipo de herramientas.

3.1. Competencias

Desde el punto de vista competencial la actividad debía estar relacionada con competencias transversales como *Razonamiento Crítico*, *Aprendizaje Autónomo* y *Adaptación a Nuevas Situaciones*. Utilizar la herramienta ChatGPT en la actividad nos permite evaluar éstas y otras competencias de una forma distinta e incluso más adecuada. Se pueden destacar las siguientes:

- **Pensamiento generativo:** habilidad para generar nuevas ideas y soluciones a partir de ideas y soluciones existentes.
- **Pensamiento crítico:** habilidad para evaluar ideas y soluciones, poniendo a prueba suposiciones y verificando afirmaciones.
- **Curiosidad:** Un comportamiento o atributo del carácter más que una habilidad; interés por explorar nuevas ideas y ámbitos; apertura a aprender cosas nuevas.

- **Interrogatorio Crítico:** habilidad para formular una secuencia de preguntas que guíen una conversación o una investigación, abran nuevas vías para el cuestionamiento y, al mismo tiempo, pongan a prueba los supuestos y verifiquen las respuestas.
- **Conocimiento del ámbito:** comprender el ámbito de la investigación es fundamental para orientar las preguntas y evaluar las respuestas.

Cuando sea necesario trabajar una competencia específica en un nivel educativo concreto, el uso de ChatGPT debería contar con diferentes mecanismos de control con el fin de no pervertir el fin de la actividad.

3.2. Elaboración de Preguntas Estructuradas

La forma en que se formula una pregunta es un aspecto de vital importancia en el proceso de obtención de resultados óptimos mediante el empleo de modelos. De esta manera, resulta imperativo establecer un diseño apropiado de las preguntas con el objetivo de maximizar la eficacia del conocimiento derivado de dichos modelos. En consecuencia, la calidad del diseño interrogatorio se erige como un factor determinante en la obtención de resultados satisfactorios. La siguiente lista enumera consejos sobre la elaboración de preguntas a modelos generadores de lenguaje.

- **Lenguaje:** Utilizar lenguaje sencillo y directo y explícito, evitando afirmaciones genéricas.
- **Variaciones:** Experimentar con múltiples variaciones de las instrucciones para obtener diversidad en los resultados.
- **Contexto:** Proporcionar contexto al modelo conseguirá poder concretar sus respuestas.
- **Razonamiento paso a paso.** Se obtendrán mejores resultados si el modelo es guiado a descomponer los problemas en subproblemas.

4. Desafíos

Los modelos generadores de lenguaje están entrenados para tareas específicas, por lo que es importante adaptar el uso del modelo a los objetivos de enseñanza del curso. Esto implica ajustar el modelo o su uso mediante las instrucciones necesarias para que sea capaz de responder preguntas específicas y proporcionar información precisa y relevante (sección 4.1).

Una limitación de ChatGPT es que puede producir respuestas aparentemente creíbles pero incorrectas. Este fenómeno, conocido como *efecto alucinación* [4], es un problema común entre muchos modelos de generación del lenguaje natural. De esta forma, es necesario un análisis riguroso de las respuestas ya que estos modelos pueden subir el fenómeno de la alucinación.

Otros aspectos como la accesibilidad del mismo para todos los estudiantes en las mismas condiciones y la integridad académica de las tareas realizadas también deben ser consideradas de forma adecuada. Por último, también es necesario considerar aspectos éticos dado que los modelos GPT-3 se han entrenado con un corpus de texto de volumen masivo, pero cuando se obtienen respuestas no son capaces de hacer referencia a qué fuentes se han utilizado para elaborar la respuesta, lo que supone una pérdida de autoría y atribución [3].

4.1. Método PEACE

El método PEACE (Preparación y planificación, Entrevista, Análisis de la información, Conclusión, Evaluación) [5] es una herramienta para ayudar a los investigadores a organizar y llevar a cabo interrogatorios. Se plantea la adaptación de este método con el fin de analizar las respuestas generadas por los modelos y evaluar su precisión y consistencia con los hechos conocidos. Este método se divide en:

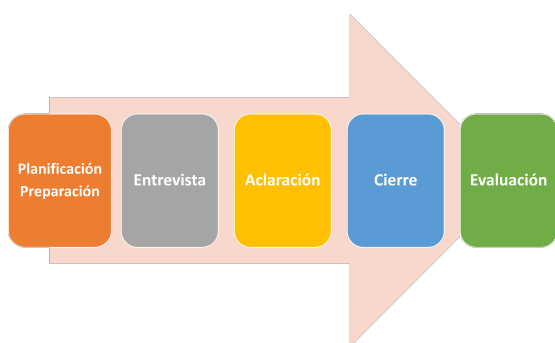


Figura 2: Etapas del Método Peace

1. Preparación y planificación: Definir una estrategia para el interrogatorio, desarrollar preguntas adecuadas, establecer un orden para las preguntas, y preparar los materiales necesarios para el interrogatorio.
2. Entrevista: Realizar las preguntas adecuadas.
3. Análisis de la información: Implica analizar la información obtenida durante el interrogatorio para identificar patrones y descubrir nuevas preguntas.
4. Conclusión: Implica obtener una conclusión sobre el interrogatorio basada en las respuestas.
5. Evaluación: Asegurar el logro de la conclusión correcta y determinar si hay nuevas preguntas que deben hacerse.

La utilización de un método como PEACE permitiría incluso mejorar la calidad de los datos de entrada utilizados para entrenar los modelos generadores de lenguaje, al evaluar y mejorar la calidad de las preguntas y respuestas utilizadas en el entrenamiento.

5. Conclusiones

El resultado tanto desde el punto de vista de la participación como de las conclusiones obtenidas ha sido más que satisfactorio lo que nos empuja a pensar que este tipo de herramientas, aun con sus inconvenientes, pueden ser integradas en las actividades docentes con el objetivo de aprovechar la accesibilidad al conocimiento, permitiendo un mejor desarrollo de competencias como el pensamiento crítico y la consulta de la información entre otras.

A partir de este punto se presentan desafíos que van desde la capacitación tanto del profesorado como de los estudiantes para un uso correcto y ético de la herramienta, en la que se posibilite el máximo desarrollo de competencias como el pensamiento crítico y la capacidad de realización de las preguntas adecuadas para conseguir un objetivo.

Referencias

- [1] Luciano Floridi y Massimo Chiriatti. GPT-3: Its nature, scope, limits, and consequences. *Minds and Machines*, 30:681–694, 2020.
- [2] Carlos Lopezosa y Lluís Codina. ChatGPT y software CAQDAS para el análisis cualitativo de entrevistas: pasos para combinar la inteligencia artificial de OpenAI con ATLAS.ti, Nvivo y MAXQDA. Barcelona: Universitat Pompeu Fabra. Departament de Comunicació, 2023. 94 p. (Serie Editorial DigiDoc. DigiDoc Reports). 2023.
- [3] Martin Miernicki e Irene Ng. Artificial intelligence and moral rights. *AI & Society*, 36:319–329, 2021.
- [4] Yiqiu Shen, Laura Heacock, Jonathan Elias, Keith D Hentel, Beatriu Reig, George Shih y Linda Moy. ChatGPT and other large language models are double-edged swords. *Radiology*, 307(2), 2023.
- [5] Brent Snook, Joseph Eastwood y W Todd Barron. The next stage in the evolution of interrogations: The PEACE model. *Canadian Criminal Law Review*, 18(2):219, 2014.
- [6] Teo Susnjak. ChatGPT: The end of online exam integrity? *arXiv preprint arXiv:2212.09292*, 2022.