



IA Generativa. Un caso de uso para el aprendizaje de Estadística en el Grado de Ciencia de Datos

Generative AI. A use case for learning Statistics in the Data Science Degree

Rosa María Alcover Arándiga^a, Vicente Chirivella González^b, Elena Vázquez Barrachina^c, Bernardo José Richart Solá^d

^aUniversitat Politècnica de València, ralcover@eio.upv.es , ^bUniversitat Politècnica de València, vchirive@eio.upv.es , ^cUniversitat Politècnica de València, evazquez@eio.upv.es  y ^dUniversitat Politècnica de València, brichart@eio.upv.es 

How to cite: Alcover Arándiga, R.M.; Chirivella González, V.; Vázquez Barrachina, E. y Richart Solá, B.J. (2024). IA Generativa. Un caso de uso para el aprendizaje de Estadística en el Grado de Ciencia de Datos. En libro de actas: *X Congreso de Innovación Educativa y Docencia en Red*. Valencia, 11 - 12 de julio de 2024.
Doi: <https://doi.org/10.4995/INRED2024.2024.18436>

Abstract

Generative artificial intelligence (Gen AI) is proposed as a work tool that every engineer should know and use. Although this is true, we have not yet reached the point where an Gen AI is capable of correctly analyse a problem based on statistics. In this document we make the choice of the most reasonable free Gen AI to use in the subject of Data Science, Statistical Models for Decision Making II, and propose its use for learning and evaluating certain aspects of this subject. As a result of a pilot test carried out, it is known that students already use the Gen AI, but have not established critical thinking about the results it offers.

Keywords: *generative artificial intelligence; AI; Gen AI; ANOVA; multiple linear regression; logistic regression; statistical analysis; interpretation of results; interactive learning; formative assessment*

Resumen

La inteligencia artificial generativa (IAG) se propone como una herramienta de trabajo que todo ingeniero debe conocer y utilizar. Aunque esto sea cierto, todavía no hemos llegado a que una IAG sea capaz de analizar de forma correcta un problema con base estadística. En este documento realizamos la elección de la IAG gratuita más razonable para utilizarla en la asignatura de Ciencia de Datos, Modelos Estadísticos para la Toma de Decisiones II, y proponemos su uso para el aprendizaje y la evaluación de ciertos aspectos de la materia. Como resultado de una prueba piloto realizada, se tiene constancia de que los alumnos ya utilizan las IAG, pero no han establecido un pensamiento crítico sobre los resultados que ofrece.

Palabras clave: *inteligencia artificial generativa; IA; IAG; ANOVA; regresión lineal múltiple; regresión logística; análisis estadístico; interpretación de resultados; aprendizaje interactivo; evaluación formativa*

Introducción

La irrupción de la Inteligencia Artificial Generativa (IAG) (Van der Zant et al., 2013) ha provocado numerosas reacciones, en ocasiones marcadamente contradictorias y divididas, respecto a sus riesgos (Chomsky et al., 2023) y bondades (Lee, 2023). Las herramientas generativas están cambiando la forma de abordar el aprendizaje y la generación de conocimiento, tanto por parte de los profesores como de los estudiantes, y plantean importantes retos para su correcta implantación en el ámbito educativo. Todo ello apunta a la necesidad de un replanteamiento de los actuales métodos de enseñanza y aprendizaje, con un enfoque pedagógico novedoso que pueda incorporar eficazmente las innovaciones impulsadas por la inteligencia artificial (Tlili, 2023).

En este sentido, el reciente documento *La inteligencia artificial generativa en la docencia universitaria* (CRUE, 2024), elaborado a partir de la Conferencia de Rectoras y Rectores de las Universidades Españolas (CRUE) señala que las IAG tienen un “potencial para apoyar y mejorar las dinámicas de enseñanza-aprendizaje, personalizando el contenido y los métodos de enseñanza, facilitando el aprendizaje colaborativo y proporcionando recursos para el aprendizaje autodirigido”. Los trabajos de García-Peñalvo (2023) y Llorens-Largo (2019) también destacan la importancia que pueden llegar a tener las herramientas de IAG en el ámbito educativo.

Conscientes de este potencial, diferentes voces en el ámbito universitario intentan motivar y promover la aplicación de la IAG en las asignaturas, sobre todo aquellas voces con un perfil más tecnológico. Algunos comentarios al respecto se encuentran en los trabajos de Sullivan et al. (2023), Sabzalieva y Valentini (2023) y García-Peñalvo (2024). Como ejemplo en la Universidad Politécnica de València (UPV) destaca la Escuela Técnica Superior de Ingeniería Informática (ETSINF UPV, 2023), familiarizada por su propia naturaleza a trabajar con IA.

En este contexto en el que se avecinan cambios, se ha desarrollado la innovación que se presenta en este trabajo. En líneas generales, el objetivo se centra en mejorar el aprendizaje de los alumnos mediante el diseño de nuevas actividades que dinamicen la docencia y permitan conseguir los resultados de aprendizaje utilizando soluciones de IAG. Más concretamente, la presente innovación se basa en el diseño de una serie de actividades que utilizan IAG como forma de aprendizaje y evaluación de los conocimientos de los alumnos, y se ha aplicado en la asignatura Métodos Estadísticos en la Toma de Decisiones II (METII), del Grado de Ciencia de Datos (GCD) impartido por la ETSINF de la UPV.

Se ha elegido METII para llevar a cabo la innovación, dado que el número de alumnos matriculados (118 alumnos) no es ni muy elevado ni muy bajo, y el número de profesores implicados (3 profesores) y su dedicación (tiempo completo) ha favorecido la coordinación en el diseño de las actividades didácticas propuestas.

No obstante, la innovación presentada podría aplicarse directamente a otras asignaturas, tales como Análisis avanzado de datos en Ingeniería Informática, asignatura optativa de 4,5 créditos del Grado de Ingeniería Informática, pues ambas comparten una parte importante de su temario.

Hasta relativamente poco tiempo, el análisis de un problema con el apoyo de la Estadística tenía (entre otros) tres pasos claramente diferenciados: definición del problema y elección del modelo, realización de cálculos, e interpretación de resultados (Peña, 2008). El uso de programas estadísticos se centraba en la parte de la realización de los cálculos, quedando la elección del modelo y la interpretación de los resultados en manos de una persona con los conocimientos adecuados. La importancia de la asignatura METII ante el alumnado radicaba en que no era posible elegir de forma automática el modelo, ni interpretar los resultados. Hasta ahora.

Con la llegada de la Inteligencia Artificial Generativa (IAG) los tres pasos se encadenan y se resuelven sin la participación, en principio, de un ser humano. O al menos, de un ser humano con conocimientos estadísticos ...

Esta es la percepción que tal vez pueden tener muchos sobre el uso de la IAG en el análisis de problemas con base estadística. Pero todavía no estamos allí, o al menos, no en la versión gratuita de una IAG. No todas la IAG pueden determinar el mejor modelo, generar el código correspondiente en R (The R Project for Statistical Computing, 2024) o Python (Python Software Foundation, 2024), ejecutarlo y recoger los resultados para interpretarlos. Y menos aún, realizar correctamente el análisis completo de principio a fin. Por tanto, mientras llegue este día, será necesario determinar de qué forma es posible utilizar la IAG para favorecer específicamente el aprendizaje de las técnicas estadísticas y evaluar los conocimientos de los alumnos de METII. Desde esta perspectiva se ha propuesto el conjunto de actividades de innovación que presentamos en este trabajo.

La estructura del trabajo es la siguiente. Tras la introducción, en el apartado 1 se abordan los objetivos de la innovación propuesta en METII. El apartado 2 presenta el desarrollo de la innovación, tanto la valoración de las IAG gratuitas en problemas de diferente naturaleza como el detalle de las nuevas actividades de aprendizaje y evaluación que se han incorporado este curso. El apartado 3 incluye los resultados obtenidos y su valoración y, finalmente, en el apartado 4 se recogen las principales conclusiones y algunas reflexiones al respecto.

1. Objetivos

El principal objetivo que nos planteamos con la innovación llevada a cabo durante el curso 2023-2024 en la asignatura METII ha sido el de incorporar actividades que permitan al estudiante valorar su conocimiento en la asignatura, detectando errores y carencias “a tiempo” y mejorar con ello su aprendizaje utilizando herramientas novedosas de inteligencia artificial.

Este objetivo general se puede descomponer en los siguientes:

1. Determinar las capacidades de las IAG gratuitas con mayor entidad en cuanto a la solución de problemas planteados en la asignatura y su interactividad con el alumno.
2. Plantear ejercicios que puedan ser resueltos mediante IAG, motivando el estudio y trabajo del estudiante y fomentando el aprendizaje de las herramientas estadísticas propuestas en la asignatura.
3. Plantear nuevas actividades de evaluación cuyo principal objetivo es el aprendizaje del alumno.

2. Desarrollo de la innovación

Se ha trabajado con los 118 alumnos matriculados durante el curso 2023-2024 en la asignatura METII. Esta asignatura troncal de 6 créditos pertenece a la materia Estadística y se imparte en el primer semestre de segundo curso del GCD de la UPV. La asignatura tiene dos grupos de matrícula, uno en horario de mañana

y el otro de tarde. Se imparte por tres profesores, y los tres han participado en el desarrollo de la innovación propuesta.

La explicación de la innovación presentada se ha estructurado de la siguiente manera. En el primer punto de este apartado se analizan las principales IAG gratuitas con el fin de determinar cuál de ellas es la más adecuada en cuanto a su capacidad para resolver el tipo de problemas que se plantean en la asignatura METII.

En el segundo apartado se detallan las actividades de aprendizaje y evaluación que se han incorporado en la docencia de la asignatura utilizando la IAG seleccionada.

2.1 Determinación de las capacidades y elección de la IA Generativa

Siguiendo el orden de los objetivos planteados en el apartado 1, en primer lugar, hay que determinar las capacidades de la IAG y seleccionar aquella que por ser más adecuada va a utilizarse en las nuevas actividades planteadas.

Se han seleccionado tres IAG gratuitas, ChatGPT, Gemini y Copilot. Simplificando su descripción se trataría de un chatbot, esto es, un programa informático diseñado para simular una conversación con seres humanos a través de interfaces de chat o mensajería, estimulado a partir de la IAG.

Aunque existen numerosas comparaciones generales entre las tres IAG referidas (ZDNET, 2024), es necesario centrarse en sus capacidades relacionadas con las principales herramientas estadísticas de la asignatura: ANOVA, Regresión lineal múltiple y Regresión logística. Para ello se van a plantear problemas-tipo de cada parte, al menos tres problemas-tipo, con las correspondientes tablas y resultados numéricos, para que las tres IAG seleccionadas contesten a las mismas preguntas en las mismas condiciones. Las preguntas realizadas se han agrupado, por su naturaleza, en cuatro ámbitos:

1. Propuesta del modelo. Se plantea un problema y se solicita a la IAG que escoja una herramienta estadística para su análisis, y que plantee el correspondiente modelo.
2. Estimación del modelo. Para el problema y modelo planteado, se solicita la generación de código en R (o Python) para estimar el modelo correspondiente.
3. Interpretación de resultados sencillos. Se solicita a la IAG que interprete resultados que puedan considerarse como conceptualmente simples: la significación de efectos o parámetros, la adecuación del modelo, el coeficiente de determinación y la predicción de la variable explicada.
4. Interpretación de resultados complejos. Se solicita a la IAG que interprete resultados que tengan una mayor complejidad conceptual: el significado de un efecto o de un parámetro, el significado de una interacción entre factores, la prueba de hipótesis para un conjunto de parámetros y la determinación de la condición operativa óptima.

En la Tabla 1 se recoge el valor SI en aquellos casos en los que la IAG responde correctamente a las solicitudes planteadas; NO para los casos en los que la IAG responde de manera incorrecta; SI(*) en los casos en los que suele responder bien, pero comete errores en ciertas cuestiones y NO(*) para los casos en los que responde mal, pero en contadas ocasiones ofrece una respuesta correcta.

Los resultados recopilados en la Tabla 1 no señalan a un claro ganador, puesto que Copilot y ChatGPT están parejos en cuanto al número de los SI/NO cosechados, pero sí hay un perdedor, que sería Gemini, que en su versión actual no es capaz de interpretar resultados complejos al nivel de los primeros. No obstante, existen ciertos comportamientos comunes que podemos señalar, y unos asteriscos que indican que la IAG no termina de dar una respuesta correcta o completa a ciertas cuestiones:

- Las IAG suelen señalar el modelo estadístico más adecuado para el análisis de una variable (ANOVA o Regresión). Sin embargo, no pueden proponer un modelo correcto cuando alguna de las variables explicativas es cualitativa. Tan sólo Gemini advierte de este hecho al usuario y apunta a que el modelo propuesto puede que no sea correcto, sin que ofrezca el modelo correcto. Por lo tanto, es necesario proporcionar un modelo correcto y detallado cuando existan variables cualitativas si se quiere obtener una respuesta razonable.
- La IAG proporciona un código R que no siempre funciona bien a la primera, o incluso que es incorrecto si hay variables cualitativas entre las explicativas, pese a que hayan sido identificadas como tales en las instrucciones.
- Las IAG interpretan de forma correcta conceptos simples, sin embargo, suelen fallar cuando interpretan un concepto que es complejo e involucra a varios elementos del problema, sobre todo si se plantea en lenguaje natural. Además, es posible que se niegue a responder si la pregunta atenta contra alguna de sus directrices internas.

Tabla 1. Capacidades de las IAG gratuitas

MODELO	ASPECTO	IAG		
		CHATGPT	GEMINI	COPILOT
ANOVA	Herramientas	SI	NO(*)	SI
	Modelo	SI	NO	SI
	Estimación del modelo	SI	SI	SI
	Interpretación de resultados sencillos	SI	SI	SI
	Interpretación de resultados complejos	SI(*)	NO	SI(*)
REGRESIÓN LINEAL MÚLTIPLE	Herramienta	SI	SI	SI
	Modelo	NO (*)	NO(*)	NO (*)
	Estimación del modelo	SI	SI	SI
REGRESIÓN LOGÍSTICA	Interpretación de resultados sencillos	SI	SI	SI
	Interpretación de resultados complejos	SI(*)	NO	SI(*)
	Modelo	SI	SI	SI
	Estimación del modelo	NO (*)	NO(*)	NO(*)
REGRESIÓN LOGÍSTICA	Estimación del modelo	SI	SI	SI
	Interpretación de resultados sencillos	SI	SI	SI
	Interpretación de resultados complejos	SI(*)	NO	SI(*)

Atendiendo a estas consideraciones, se escoge Copilot. Esta IAG suele ofrecer los mismos resultados que ChatGPT, y es mejor que Gemini. Además, y en teoría, esta opción debería ser la mejor al estar basado en GPT-4, y los alumnos ya disponen de Copilot con su cuenta upv.edu.es.

2.2 Actividades de Aprendizaje-Evaluación

El segundo objetivo de la innovación propuesta es que los alumnos adquieran destreza en la solución de problemas estadísticos, y para ello se plantea la propuesta de ejercicios, basados en problemas reales, a realizar durante las sesiones de seminario. La resolución de estos será presencial, en grupo, y con una duración máxima de dos horas, aunque también habrá ejercicios para realizar fuera del aula.

En primer lugar, los alumnos analizan un problema planteado por el profesor, interpretan el problema, proponen el modelo, lo estiman (proporcionando las tablas y los gráficos correspondientes), y responden a ciertas preguntas clave relacionadas con el enunciado. Los alumnos dispondrán de la resolución del problema al terminar su análisis, para autoevaluarse, detectando interpretaciones, cálculos o conceptos erróneos, mejorando su conocimiento y favoreciendo su aprendizaje en la materia.

Una vez adquiridos los conocimientos, los alumnos resuelven las preguntas clave de estos mismos ejercicios con la ayuda de la IAG. Preparan unas instrucciones claras y completas con el problema que se pretende resolver, detallan las preguntas que deben responderse, y seleccionan las tablas necesarias con las estimaciones del modelo.

Planteada la cuestión a la IAG, los alumnos toman nota de las respuestas ofrecidas a cada una de las preguntas y determinan si la respuesta ofrecida es correcta, y en caso de no serlo, el motivo por el que no lo es. Esta respuesta incorrecta puede ser debida a un mal planteamiento del problema o de la pregunta, o puede ser un error del programa. En el primer caso se plantea de nuevo el problema y las preguntas. En el segundo, se solicita una nueva respuesta, o se interroga interactivamente a la IAG sobre los motivos de la respuesta ofrecida, o bien se la confronta con la respuesta correcta, para ver su argumentación. La IAG ofrece secuencialmente múltiples respuestas, muchas veces contrarias entre sí, por lo que es imprescindible repreguntar para ver cómo cambia la respuesta.



Departamento de Estadística e Investigación Operativa Aplicadas y Calidad
Modelos Estadísticos para la Toma de Decisiones II
Grado en Ciencia de Datos ETSINF - UPV
Examen Final 09/02/2024



P2.- Se desea explicar la probabilidad de solicitar una cita médica a través de internet, para lo cual el INE realizó una encuesta en octubre del año 2017. Con sus resultados, se plantea un modelo de primer orden que explica si se solicitó la cita (CITAINTERNET, valor 1 si se solicitó por internet y 0 en otro caso), a partir de su sexo, edad y estado civil de la persona entrevistada, y de si esta persona es un enfermo crónico. Se escoge como referencia los valores mayores de las variables (según la siguiente tabla), y se restan 30 años a la edad del entrevistado. Justifica todas tus respuestas.

$$\ln(p/(1-p)) = \beta_0 + \beta_1(EDAD-30) + \beta_{2,1}SEXO=1 + \beta_{3,1}ECIVIL=1 + \beta_{3,2}ECIVIL=2 + \beta_{3,3}ECIVIL=3 + \beta_{3,4}ECIVIL=4 + \beta_4CRONICO=1$$

sexo del entrevistado SEXO	1 hombre 2 mujer	estado civil ECIVIL	1 casado/a 2 soltero/a
es enfermo crónico CRONICO	1 sí 2 no		3 viudo/a 4 separado/a 5 divorciado/a
edad del entrevistado EDAD	años		

Preguntado a Copilot sobre la interpretación de los efectos de las variables explicativas según el signo de su estimación, y ofreciendo las mismas tablas de este enunciado, se ha obtenido los siguientes resultados. Determina si las afirmaciones realizadas por el programa son correctas o no, justificando el motivo por el que lo son, e introduce las correcciones necesarias en las interpretaciones dadas. (1,0p)

1. ****CONSTANTE (Intercepto):****
 - Estimado: 0.00903592
 - Interpretación: Este es el término constante o intercepto del modelo. Representa la probabilidad estimada de solicitar una cita médica por internet cuando todas las variables explicativas son cero. En este caso, es muy cercano a cero.
2. ****EDAD-30:****
 - Estimado: -0.0354063
 - Interpretación: Por cada año que una persona disminuye su edad (ajustada restando 30 años), la probabilidad de solicitar una cita médica por internet disminuye en 0.0354. Dado que el coeficiente es negativo, se espera que a medida que la edad disminuya, la probabilidad de solicitar una cita en línea también disminuirá.
3. ****SEXO=1 (Hombre):****
 - Estimado: -0.219266
 - Interpretación: Comparado con las mujeres (referencia al ser igual a 2), los hombres tienen una probabilidad estimada 0.2192 menor de solicitar una cita médica por internet. Dado que el coeficiente es negativo, se espera que las mujeres tengan una probabilidad más alta que los hombres de solicitar citas médicas en línea.
4. ****ECIVIL=1 (Casado):****
 - Estimado: -0.262104
 - Interpretación: Comparado con las personas solteras (referencia al ser igual a 2), las personas casadas tienen una probabilidad estimada 0.2621 menor de solicitar una cita médica por internet. Dado que el coeficiente es negativo, se espera que las personas solteras tengan una probabilidad más alta que las personas casadas de solicitar citas médicas en línea.
5. ****ECIVIL=2 (Viudo):****
 - Estimado: -0.701354
 - Interpretación: Comparado con las personas solteras (referencia al ser igual a 2), las personas viudas tienen una probabilidad estimada 0.7014 menor de solicitar una cita médica por internet. Dado que el coeficiente es negativo, se espera que las personas solteras tengan una probabilidad más alta que las personas viudas de solicitar citas médicas en línea.
6. ****ECIVIL=3 (Separado):****
 - Estimado: -0.791261
 - Interpretación: Comparado con las personas solteras (referencia al ser igual a 2), las personas separadas tienen una probabilidad estimada 0.7913 menor de solicitar una cita médica por internet. Dado que el coeficiente es negativo, se espera que las personas solteras tengan una probabilidad más alta que las personas separadas de solicitar citas médicas en línea.

Fig. 1. *Uso de la IAG en la evaluación de respuesta abierta*

Esta interacción ofrecida por la IAG puede facilitar y ayudar a mantener el interés, la motivación y la comprensión de conceptos, así como mejorar el desarrollo de ciertas habilidades del alumno en su proceso de aprendizaje. En ocasiones, discriminar si la respuesta proporcionada por la IAG es o no correcta puede no ser trivial para el alumno. En este sentido, para evaluar las respuestas de la IAG se requieren conocimientos, por ello el alumno deberá revisar aquellas partes de la materia que le generen duda ante las respuestas de la IAG. El estudiante además de autoevaluarse, detectando conceptos mal aprendidos, procedimientos o razonamientos incorrectos, con este tipo de actividades trabaja y fomenta el pensamiento crítico, muy importante en la sociedad actual.

Finalmente, y en relación con el objetivo 3 de la innovación presentada en este trabajo, la evaluación de los conocimientos estadísticos en METII puede realizarse a través del trabajo académico de la asignatura y a partir de la prueba escrita de respuesta abierta, ambas con la asistencia de la IAG.

En el caso del trabajo académico de la asignatura, se procede con un nuevo problema, cuya solución sea aportada por la IAG y esté supervisada por el alumno. En teoría, desde la propuesta del modelo, pasando por la indicación de la forma de proceder con el programa estadístico (o la generación del código) y terminando con la interpretación de resultados, el alumno debe valorar y corregir, en su caso, la respuesta de la IAG. Se trata de un proceso complejo, que motiva al alumno al tratarse de un desafío ante la IAG y aunque este se lo pueda tomar como un juego, resulta muy enriquecedor para su aprendizaje. Además, toda esta interacción de preguntas y respuestas con la IAG se realiza de forma inmediata.

En cuanto a la prueba escrita de respuesta abierta, se plantea un problema, con ciertas preguntas clave cuya respuesta se solicita a la IAG. Los alumnos deben determinar si la interpretación dada por la IAG es correcta, y en caso de no serlo deben señalar y justificar los errores existentes y aportar la interpretación correcta. Se trataría de una actividad similar a las ya realizadas en el aula, pero más breve, al tratarse de una prueba de evaluación, y sin la interacción que permite el ordenador. En la Fig. 1 se muestra un ejemplo de los utilizados en la asignatura.

3. Resultados obtenidos

Para valorar los efectos de la incorporación de la IAG sobre el aprendizaje de los alumnos de METII, se realizó una prueba piloto durante el curso académico 2023-2024 en algunas sesiones de seminario y en la evaluación final de la asignatura (prueba escrita de respuesta abierta).

La primera parte de la prueba piloto abarcó algunas de las sesiones de seminario de la asignatura. Respecto a las actividades llevadas a cabo durante estas sesiones, se plantearon diferentes ejercicios prácticos a partir de enunciados disponibles para los alumnos. Cada ejercicio práctico o problema implicaba crear un fichero de instrucciones para la IAG y plantear preguntas concretas relacionadas con parámetros del modelo, para contrastar las respuestas ofrecidas por la misma y así, valorar su veracidad y utilidad. Se realizaron actividades de esta naturaleza tanto en la parte del temario correspondiente a modelos de regresión lineal múltiple como en la de regresión logística.

Lamentablemente, la mayor parte de estas actividades no fue “totalmente exitosa”, en el sentido de que pocos grupos de alumnos utilizaron abiertamente la IAG en el aula (en torno al 60% del grupo de mañana y 40% en el grupo de tarde). En general, el grupo de mañana obtuvo mejores resultados con la IAG.

Por parte del profesorado se esperaba una mayor implicación del alumnado en estas actividades. El horario de las sesiones de seminario en ambos grupos de la asignatura (viernes a última hora) quizá no era muy favorable para llevarlas a cabo, pero por ello, se confió que esta sería una buena medida para estimular la participación y el aprendizaje activo, sobre todo durante franjas horarias en las que los alumnos pudieran estar más cansados.

La segunda parte de la prueba piloto consistió en la evaluación de la asignatura mediante preguntas de respuesta abierta. Se decidió considerar este tipo de evaluación en la prueba piloto por el siguiente motivo. Durante la corrección por parte del profesorado, tanto del trabajo académico de asignatura como de la primera de las evaluaciones de respuesta abierta realizada a los estudiantes, se detectó un cierto número de respuestas incorrectas, casualmente con una misma redacción, una respuesta incorrecta ofrecida habitualmente por Copilot y ChatGPT.

Según esto, parece razonable pensar que algunos alumnos habitualmente “ausentes” en el aula, utilizan la IAG para estudiar y, lamentablemente, aprenden con las respuestas incorrectas generadas por el programa (que se denominarán en adelante como respuestas “incorrectas creativas”). Por ello, estos alumnos

proporcionan a las preguntas planteadas una respuesta bastante uniforme y muy diferente de las respuestas trabajadas y proporcionadas por los alumnos presentes en el aula. Mientras que en los años previos a la IAG era inusual este tipo de respuestas, actualmente es bastante más probable.

Por este motivo, la segunda parte de la prueba piloto consistió en que los alumnos valoraran la veracidad de las interpretaciones ofrecidas por Copilot, utilizando la segunda evaluación de respuesta abierta realizada en la asignatura. De esta forma sería posible cuantificar el porcentaje de alumnos que ofrecen respuestas “incorrectas creativas”.

En cuanto a los resultados obtenidos en las dos evaluaciones (prueba escrita de respuesta abierta) se ha calificado la respuesta de los alumnos ante las interpretaciones ofrecidas por Copilot en diferentes problemas como correcta, incorrecta, incorrecta creativa y en blanco, que corresponden, respectivamente a una respuesta correcta, una respuesta incorrecta, una respuesta incorrecta de las ofrecidas habitualmente por la IAG, o que no hubo respuesta por parte del alumno.

Se han comparado los resultados de las dos evaluaciones de respuesta abierta realizadas, obteniéndose los siguientes resultados:

Tabla 2. Tipo de respuesta

Evaluación de respuesta abierta	CORRECTA	INCORRECTA	INCORRECTA CREATIVA	EN BLANCO
Primera	56,6%	26,5%	8,9%	8,0%
Segunda	54,5%	20,2%	9,1%	16,2%

De estos resultados se constata que un 9% de los alumnos evaluados aprenden con las interpretaciones ofrecidas por la IAG y lo transcriben directamente. Estos alumnos no se plantean que la respuesta proporcionada por la IAG pudiera ser incorrecta, y en la valoración de la interpretación dada por la IAG se limitan a una reinterpretación simple del resultado en el ámbito del problema planteado.

Este resultado obtenido (9%) pone de manifiesto que el uso de la IAG en METII, sin la correspondiente valoración crítica del alumno en las respuestas proporcionadas, produce el efecto contrario al deseado en el aprendizaje de los contenidos de la asignatura.

4. Conclusiones

En este trabajo se ha presentado la innovación llevada a cabo en la asignatura METII, de 6 créditos y carácter obligatorio, impartida en el segundo curso del GCD de la UPV. La innovación ha consistido en la incorporación de la IAG gratuita en diferentes actividades docentes. Todas estas actividades presentan un propósito común facilitar, favorecer y mejorar los resultados de aprendizaje planteados en la asignatura para los estudiantes. Para ello, se ha analizado la utilidad de las principales IAG gratuitas existentes en la actualidad, ChatGPT, Gemini y Copilot, en el contexto de las herramientas estadísticas presentadas en la asignatura. Como resultado, la mejor propuesta es Copilot para las actividades de aprendizaje-evaluación en la misma.

Tras las actividades llevadas a cabo, se ha determinado que los análisis realizados por las IAG pueden ofrecer conclusiones incorrectas cuando el concepto tratado es complejo. También se ha detectado que las IAG no son capaces de manejar variables cualitativas en el contexto del ANOVA y de los modelos de regresión. No proponen modelos adecuados, ni generan el código de R correcto, ni son capaces de interpretar correctamente los resultados, aunque se les ofrezca por otros medios.

El análisis cuantitativo ha revelado que algunos alumnos (9%) prefieren “aprender” interactuando con una IAG a utilizar el material docente proporcionado en la asignatura, y que, además, otorgan un mayor valor a sus respuestas, sin realizar ninguna valoración crítica al respecto. Dado que en muchas ocasiones las respuestas son incorrectas, este es un descubrimiento muy preocupante. Será conveniente insistir en actividades en el aula que refuercen el espíritu y pensamiento crítico de los estudiantes.

Mientras no mejoren las respuestas ofrecidas por la IAG en el ámbito de la asignatura, se va a reducir su uso a determinar la veracidad de los resultados que ofrece, pero no a la selección y estimación de modelos. Por tanto, se propone usarla solo para el análisis, concretamente, para responder a cuestiones sencillas de interpretación en los problemas planteados en las sesiones de práctica de aula y de laboratorio. También se prevé utilizarla para valorar la veracidad de afirmaciones sobre los resultados proporcionados, en algún apartado de la evaluación con respuesta abierta.

Finalmente, y como ya se ha mostrado, estas herramientas generativas pueden parecer peligrosas por los riesgos que su mal uso pudiera conllevar (falta de creatividad, vulneración de los derechos de autoría, dependencia de la tecnología, ...). Sin embargo, “han venido para quedarse” y como docentes no se puede obviar los aspectos positivos que pueden ofrecer (aprendizaje personalizado, potenciador del pensamiento crítico, mejora de la accesibilidad de la información, ...). Ante este nuevo escenario, es necesaria una profunda reflexión sobre cómo pueden incorporarse a nuestras asignaturas con el fin de que el alumno sepa utilizarlas correctamente, con espíritu crítico y en beneficio de su propio aprendizaje. También se tendrá que capacitar al profesorado para un uso correcto y ético de estas herramientas de forma que se pueda extraer todo su potencial en el proceso de enseñanza/aprendizaje.

Indudablemente, se trata de un gran desafío, pero los docentes no podemos perder esta oportunidad, pues está en juego algo muy importante. El futuro está condicionado por la educación. En este sentido, la propuesta va en la línea de avanzar progresivamente y con cautela en su incorporación, a la vez que se valoran todos sus efectos en los procesos de enseñanza/aprendizaje.

Referencias

- Chomsky, N., Roberts I. y Watumull, J. (2023). The False Promise of ChatGPT. *The New York Times*. <http://bit.ly/3GycXfx>
- Conferencia de Rectoras y Rectores de las Universidades Españolas. (2024). *La inteligencia artificial generativa en la docencia universitaria* [Archivo PDF]. https://www.crue.org/wp-content/uploads/2024/03/Crue-Digitalizacion_IA-Generativa.pdf
- ETSINF UPV. (2023). *ChatGPT y otras herramientas generativas en la Educación Universitaria. Acciones a emprender*. Twitter. <https://twitter.com/etsinfupv/status/1730511741392502979>
- García-Peñalvo, F.J. (2023). *Cómo se percibe la Inteligencia Artificial en la educación tras el lanzamiento de ChatGPT*. <http://repositorio.grial.eu/handle/grial/2890>
- García-Peñalvo, F. J. (2024, enero 13). Inteligencia Artificial Generativa en la Educación Superior: Una Perspectiva de 360°. IFE Conference 2024 (IFE 2024), Monterrey, México. Grupo GRIAL. <https://doi.org/10.5281/zenodo.10499827>
- Lee, E. (2023). *Is ChatGPT a False Promise?* Berkeley Blog. <http://bit.ly/3UIHsv1>
- Llorens-Largo, F. (2019). *Las tecnologías en la educación: características deseables, efectos perversos*. <https://bit.ly/3SxO72D>
- Peña D. (2008). *Fundamentos de estadística*. Alianza Editorial.

- Python Software Foundation (06 de junio de 2024). *About*. <https://www.python.org/>
- Sabzalieva, E. y Valentini, A. (2023). *ChatGPT e inteligencia artificial en la educación superior: guía de inicio rápido*. UNESCO e Instituto Internacional de la UNESCO para la Educación Superior en América Latina y el Caribe. https://unesdoc.unesco.org/ark:/48223/pf0000385146_spa
- Sullivan, M., Kelly, A. y McLaughlan, P. (2023). ChatGPT in higher education: Considerations for academic integrity and student learning. *Journal of Applied Learning and Teaching*, 6 (1). <https://doi.org/10.37074/jalt.2023.6.1.17>.
- The R Project for Statistical Computing (06 de junio de 2024). *About R*. <https://www.r-project.org/>
- Tlili, A., Shehata, B., Adarkwah, M.A., Bozkurt, A., Hicket, D.T., Huang, R. y Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*, vol. 10, no. 1. doi: 10.1186/s40561-023-00237-x.
- Van der Zant, T., Kouw, M., & Schomaker, L. (2013). *Generative Artificial Intelligence: Philosophy and Theory of Artificial Intelligence*. In V. C. Mueller (Ed.), *Philosophy and Theory of Artificial Intelligence (Vol. 5, pp. 107-120)*. (Studies in Applied Philosophy, Epistemology and Rational Ethics; Vol. 5). Springer. https://doi.org/10.1007/978-3-642-31674-6_8
- ZDNET. (9 de febrero de 2024). *ChatGPT vs. Microsoft Copilot vs. Gemini: Which is the best AI chatbot?*. <https://www.zdnet.com/article/chatgpt-vs-microsoft-copilot-vs-gemini-which-is-the-best-ai-chatbot/>