

## Integración de una visión humanística en la asignatura de aprendizaje automático

Eduardo C. Garrido-Merchán  
Carlos Martínez de Ibarreta Zorita  
Universidad Pontificia Comillas

**Resumen:** El auge de la computación ha hecho posible experimentar con modelos de aprendizaje automático, o machine learning, de billones de parámetros, como Dalle-2 o GPT-3. Estos modelos, con complejas arquitecturas, representan lenguajes, imágenes o problemas complejos. En ocasiones, estos sistemas son capaces de solucionar problemas con mayor eficacia que la de un ser humano. En concreto, solucionan el problema del plegamiento de proteínas, el comercio de alta frecuencia, la elaboración de nuevas teorías matemáticas y físicas, la escritura de códigos de programación o el Go o al ajedrez. Ante este paradigma de imparable evolución, es imprescindible una reflexión crítica sobre las repercusiones éticas que presenta el uso de estas tecnologías. Este artículo presenta el alcance práctico del machine learning y motiva las razones por las que es fundamental que el alumno no sea soloun experto técnico de la materia, sino que también alinee sus conocimientos a su escala de valores. En concreto, se presentan dos actividades de carácter humanístico, una charla de divulgación y un debate, que los docentes de la asignatura de machine learning han llevado acabo en la facultad de Ciencias Económicas y Empresariales de la Universidad Pontificia Comillas de Madrid. Adicionalmente, presentamos los resultados de una encuesta efectuadaa nuestro alumnado sobre estas actividades.

### Introducción

El mundo está sufriendo un cambio de paradigma en lo respectivo a la tecnología que las empresas utilizan para realizar sus actividades. En concreto, el avance de la computación posibilita la implementación de metodologías basadas en el aprendizaje automático (machine learning) capaces de resolver tareas antes impensables. En particular, estos algoritmos han impactado significativamente en sectores como el financiero (Dixon, 2020), el energético (Lay, 2020) o el de la salud (Shailaja, 2018), entre otros.

Este cambio de paradigma ha hecho que, en las universidades, estén apareciendo nuevos grados y planes de estudios para formar en el conocimiento y la aplicación de estas nuevas técnicas. En concreto, las personas con formación en este campo son de los profesionales más demandados en la actualidad. En particular, en la Universidad Pontificia Comillas, dentro de la Facultad de Ciencias Económicas y Empresariales, se ha desarrollado el grado en Business Analytics, para dar respuesta a esta demanda de formación. Entre las materias más técnicas incluidas en dicho grado están tres asignaturas sobre machine learning, en las que los autores somos profesores.

Los autores somos conscientes de que es necesario añadir a los contenidos técnicos una reflexión de carácter humanístico sobre las repercusiones que el uso de esta tecnología provoca en la sociedad y hacer al alumno coparticipe de estas reflexiones.

Esto está completamente alineado con la misión de nuestra facultad: “formar líderes empresariales competentes, conscientes, críticos y compasivos, capaces de transformar el mundo” (Universidad Pontificia Comillas, 2022).

Pero, por desgracia, el temario de la asignatura es amplio y de carácter técnico, lo que resulta difícil al alumnado. Por tanto, esta tarea no es fácil de integrar en la planificación docente, corriendo el riesgo de que sea relegada.

Con el fin de solucionar estas carencias, se presenta en este trabajo un estudio sobre la realización de dos actividades complementarias a la docencia obligatoria: una conferencia de carácter divulgativo de cómo el machine learning está transformando nuestra sociedad y un debate sobre las repercusiones sociales, tanto positivas como negativas, de estos algoritmos. En particular, pensamos que es importante no solo exponer los beneficios sino también los riesgos, que los alumnos tienden a percibir distorsionados por la literatura y cine de ciencia ficción a las que están expuestos.

Nuestro objetivo es también servir de precedente para que otras asignaturas técnicas de la facultad cuyo contenido transforma a la sociedad, como por ejemplo el trading algorítmico, incorpore estas actividades para que el alumno adquiera los valores de la misión de la facultad. Pese a que nuestros planes de estudios contemplan la Ética como asignatura evaluada, creemos que queda camino hasta conseguir una visión transversal, en la cual este tipo de conocimientos se integren en cada una de las asignaturas técnicas cuyo contenido tiene alta probabilidad de transformar nuestra sociedad.

Para ilustrar la realización de estas actividades, hemos organizado este artículo de la siguiente manera: en primer lugar, se describen las razones por las que es importante efectuar una reflexión sobre las repercusiones sociales de la práctica del machine learning (en adelante ML) en la empresa. Para ello, se necesita especificar qué problemas sociales son importantes en la actualidad. Con este propósito, se desarrolla una sección que estudia la conciliación de la práctica del ML, con los problemas sociales mencionados en la encíclica *Laudato Si* (Francisco, 2015). Tras esta motivación, se presentan los detalles de las dos actividades de corte humanístico que los docentes de la asignatura de ML hemos impartido en la Universidad Pontificia Comillas. Posteriormente, mostramos un resumen del cuestionario que los alumnos de esta asignatura han cumplimentado. Este cuestionario versa sobre su percepción de las actividades humanísticas efectuadas en la asignatura de ML y la importancia que tiene la problemática del auge de la inteligencia artificial en la sociedad. Por último, terminamos el artículo exponiendo nuestras conclusiones y líneas de trabajo futuro relacionadas con esta investigación.

### **Una reflexión sobre las repercusiones del machine learning según la encíclica *Laudato Si***

En esta sección, nuestro propósito es alinear la visión propuesta en la encíclica *Laudato Si* al nuevo paradigma social fruto de la práctica de la inteligencia artificial (en adelante IA) y, en concreto, el ML. Pese a los potenciales beneficios que pueden ocasionar estas metodologías, es importante reflexionar sobre sus riesgos. Por ejemplo, el uso de estas metodologías puede repercutir negativamente al medio ambiente, distorsionar la concepción que la sociedad tiene sobre el arte, acentuar la desigualdad e incluso olvidar qué nos hace realmente humanos.

En particular, la encíclica Laudato Si (Francisco, 2015) comienza con el siguiente extracto, con el que también deseamos comenzar esta sección:

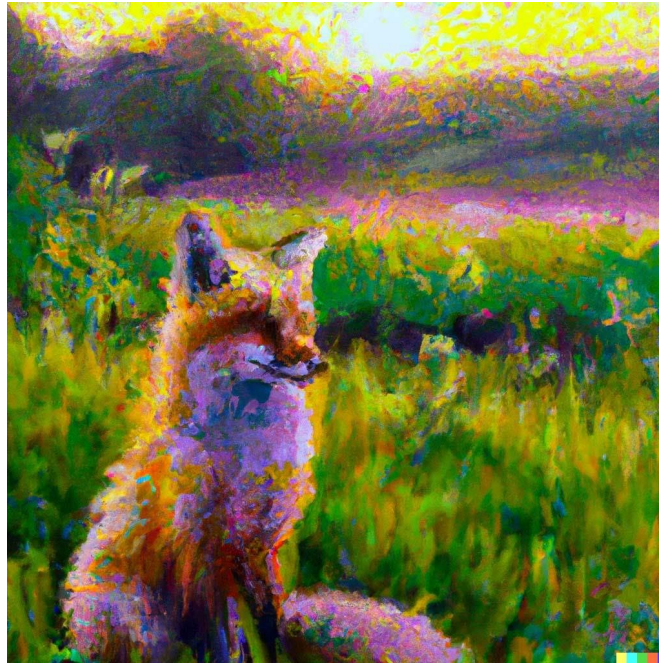
*« Laudato si ’, mi’ Signore » – « Alabado seas, mi Señor », cantaba san Francisco de Asís. En ese hermoso cántico nos recordaba que nuestra casa común es también como una hermana, con la cual compartimos la existencia, y como una madre bella que nos acoge entre sus brazos: « Alabado seas, mi Señor, por la hermana nuestra madre tierra, la cual nos sustenta, y gobierna y produce diversos frutos con coloridas flores y hierba. (1) »*

Nuestra casa común ahora va a ser compartida con la IA, que sería al fin y al cabo un bien que Dios nos ha otorgado (*Esta hermana clama por el daño que le provocamos a causa del uso irresponsable y del abuso de los bienes que Dios ha puesto en ella. (2)*). Por lo tanto, debemos reflexionar sobre las repercusiones que este hecho puede ocasionar en nuestra casa común antes de que sea demasiado tarde, por ejemplo, a causa del cambio climático.

Con respecto al cambio climático, la encíclica reconoce este problema (*Hay un consenso científico muy consistente que indica que nos encontramos ante un preocupante calentamiento del sistema climático. (23)*) y el ML lo agrava. Por ejemplo, el entrenamiento de los modelos de ML ocasiona el 10-15% de la contaminación emitida por Google (Patterson, 2021). Este hecho es grave si se generaliza para todo el tejido empresarial, no únicamente perteneciente al sector tecnológico. No obstante, existen vías para que el entrenamiento de estos modelos no provoque tanta contaminación (Candelieri, 2021). Por tanto, es importante que el alumnado sea consciente de este hecho para que, una vez adquiridos los conocimientos necesarios para entrenar algoritmos de ML, se tome en cuenta el factor climático para decidir que metodología en concreto usar.

Como ya se ha mencionado, la IA ha incrementado su aplicabilidad en los últimos años, abanderada por el auge tecnológico y el rendimiento de los modelos más innovadores basados en ML y redes neuronales (LeCun, 2015). Por ejemplo, incluso en el ámbito del arte, el modelo DALLE-2 es capaz de generar imágenes completamente innovadoras sin intervención humana a partir de un texto (Ramesh, 2022).

Figura 1. Imagen generada por el modelo DALLE-2 cuando se le introduce el texto “a painting of a fox sitting in a field at sunrise in the style of Claude Monet” sin intervención humana.



Fuente: Página web de OpenAI donde se exponen los resultados del modelo DALL-E-2  
<https://openai.com/dall-e-2/>

En consecuencia, sino se reflexiona sobre el concepto de arte, la sociedad puede incurrir en un nuevo tipo de degradación social, correspondiente a la humanización de las máquinas, capaces de efectuar cualquier tarea antes solo atribuible a seres humanos, y la deshumanización del ser humano (Haslam, 2006), cada vez más conectado a través de dispositivos que le suministran la información que debe consumir con algoritmos basados en ML. La progresiva pérdida de comunicación sin medios digitales incurre en un nuevo deterioro de la calidad de la vida humana, ya que ante todo esta es fruto de compartir experiencias con seres queridos y amistades cercanas. Este hecho sería un síntoma de lo que la encíclica cita como *“síntomas de una verdadera degradación social, de una silenciosa ruptura de los lazos de integración y de comunión social, dinámicas de los medios del mundo digital que, cuando se convierten en omnipresentes, no favorecen el desarrollo de una capacidad de vivir sabiamente, de pensar en profundidad, de amar con generosidad”*. (46). Si cabe, este hecho es aún más importante ya que como, de nuevo, cita la encíclica, *“no podremos afrontar adecuadamente la degradación ambiental si no prestamos atención a causas que tienen que ver con la degradación humana y social”*. (48).

Ante una decisión empresarial que incurre, dentro del contexto mencionado, en una humanización de las máquinas y una deshumanización del ser humano, el alumno ha reflexionado en el debate celebrado y mediante los contenidos presentados en la divulgación efectuada en la semana de la ciencia celebrada en la facultad.

Estos riesgos son aún más graves si se tiene en cuenta que la mayor parte de estos puntos oscuros quedan muchas veces enmascarados por la brillante plétora de utilidades y aplicaciones positivas que para los individuos y la sociedad tiene la generalización de estas herramientas y aplicaciones tecnológicas. Como cita la encíclica:

*“La humanidad ha ingresado en una nueva era en la que el poderío tecnológico nos pone en una encrucijada” (102)... “Pero no podemos ignorar que la energía nuclear, la biotecnología, la informática, el conocimiento de nuestro propio ADN y otras capacidades que hemos adquirido nos dan un tremendo poder. Mejor dicho, dan a quienes tienen el conocimiento, y sobre todo el poder económico para utilizarlo, un dominio impresionante sobre el conjunto de la humanidad y del mundo entero”. (104) ... “El problema fundamental es otro más profundo todavía: el modo como la humanidad de hecho ha asumido la tecnología y su desarrollo junto con un paradigma homogéneo y unidimensional” (106) ... “La especialización propia de la tecnología implica una gran dificultad para mirar el conjunto” (110).*

Pero el problema no reside únicamente en los riesgos de las aplicaciones sino también en la capacidad de entrenar un modelo de ML. En concreto, en el caso de las redes neuronales más avanzadas, esta tarea es únicamente accesible a aquellas entidades que cuenten con recursos computacionales adecuados. Si se suma que mediante el uso de estos modelos podremos crear nuevos sistemas, o incluso aumentos como pretende NeuraLink (Kulshreshth, 2019), que supondrán una modificación del estilo de vida de las personas que puedan acceder a ellos, esto generará desigualdad e inequidad planetaria. *“La inequidad no afecta sólo a individuos, sino a países enteros, y obliga a pensar en una ética de las relaciones internacionales”. (51)*

La reacción por parte de la humanidad contra la potencial desigualdad que puede provocar el ML debe ser contundente e inmediata. Bien enfocado, el ML, aplicado en la robótica, por ejemplo, puede ayudar a crear una sociedad menos dependiente de trabajos con malas condiciones, colaborando al bienestar de nuestra casa común. Por ello, tal y como se menciona en la encíclica:

*“Se vuelve indispensable crear un sistema normativo que incluya límites infranqueables y asegure la protección de los ecosistemas, antes que las nuevas formas de poder derivadas del paradigma tecnoeconómico terminen arrasando no sólo con la política sino también con la libertad y la justicia. (53)”*

Por ejemplo, iniciativas como la regulación europea que obliga a que los algoritmos usados de IA sean explicables para conceder una hipoteca se vuelven esenciales. No hacerlo sería brindar a las nuevas formas de poder (basadas en la IA), una vía para acabar con la libertad y la justicia individual. Por ejemplo, no conceder una hipoteca a una persona solo por su color de piel o por ser un activista. En consecuencia, no es válido un progreso en el ML que deje al margen la calidad de las personas a costa de su aplicación indiscriminada *(Simplemente se trata de redefinir el progreso. Un desarrollo tecnológico y económico que no deja un mundo mejor y una calidad de vida integralmente superior no puede considerarse progreso. (194))*

Finalmente, se hace crítica una reflexión sobre los límites a los que llegará la aplicación del ML y, en concreto, las redes neuronales en la sociedad. Movimientos como el transhumanismo (Bostrom, 2005) promocionan la implantación de dispositivos que usan IA en seres humanos. El propósito de estos dispositivos es aumentar nuestras capacidades, afectando a nuestra consciencia. Su uso indiscriminado podría afectar nuestro concepto del

ser humano, peligrosamente cuantitizando al ser en función de sus habilidades en vez de a su esencia, como por ejemplo en base a una estimación externa de su inteligencia. Peor aún, la apariencia externa de un sistema basado en ML podría llegar a ser indistinguible de un ser humano, si olvidamos que lo que nos hace humanos es nuestro ser, el observador de la consciencia fenoménica (Block, 1995), y no la mera inteligencia computacional, que sería una propiedad independiente (Garrido-Merchán, 2021). En consecuencia, una máquina podrá ser siempre reemplazable, pero un ser vivo será irremplazable. En palabras de Laudato Si, sustituyendo a “las cosas”, por “la IA”:

*“A la vez que podemos hacer un uso responsable de la IA, estamos llamados a reconocer que los demás seres vivos tienen un valor propio ante Dios y, por su simple existencia, lo bendicen y le dan gloria. (69)”*

### **Integrar la misión con actividades opcionales: divulgación y debate**

Una vez se ha alineado la visión de Laudato Si con el nuevo paradigma de la IA, exponemos ahora como integrar dicha visión en la misión de nuestra universidad a través de la asignatura de ML. Como se ha comentado anteriormente, los planes de estudios en asignaturas de tipo técnico tienen unos contenidos amplios y de cierta dificultad a impartir en un tiempo limitado. Eso provoca muchas veces que, a pesar de que se considere importante, la reflexión y discusión sobre los aspectos éticos y humanísticos no esté contemplada expresamente en dichos planes o que no se vea con la suficiente profundidad por falta de tiempo.

Por ese motivo, y para intentar un mejor alineamiento con la visión y valores católicos, expresados en la doctrina de la Iglesia en general y con la misión de la Universidad Pontificia Comillas en particular, se decidió organizar dos actividades complementarias y voluntarias en la asignatura: (1) un debate y (2) una conferencia sobre los aspectos positivos y negativos de las técnicas de Deep Learning. En la primera se buscaba sobre todo la mejorar la motivación y el interés a través de una actividad voluntaria y que fomenta la libertad de expresión. En la segunda, se pretendía alejar complejidad técnica para democratizar las repercusiones en la sociedad del ML y efectuar una formación integral del alumno independiente de su carrera específica. A continuación, se describen en detalle ambas actividades.

Respecto al debate, se planteó como actividad voluntaria para los alumnos, a realizar fuera del horario de clase, con una duración de dos horas. Se tomaron como modelos los torneos de debate universitario, en los cuales la Universidad Pontificia Comillas es un referente en el panorama español e internacional, contando con numerosos equipos ganadores de campeonatos y trofeos.

Se pensó que era conveniente ofrecer de manera optativa esta actividad para complementar los contenidos técnicos vistos en la asignatura de ML 1 y que los alumnos, de forma autónoma y en un contexto dialéctico fueran sacando a la luz aquellas cuestiones de carácter ético y humanístico derivadas de la implantación general de algoritmos de ML e IA en diversos ámbitos de la economía o la sociedad, como por ejemplo la justicia, las finanzas, la salud, la seguridad o la creación artística.

Los alumnos que deseaban participar debían inscribirse en grupos de cuatro personas, designando un portavoz. Los equipos fueron asignados al azar a tener que defender la postura “a favor” o “en contra” de la aplicación de dichas técnicas de ML e IA.

Los profesores actuaron como moderadores y jurados, valorando y puntuando el número de intervenciones de cada equipo, la calidad y fundamentación de las argumentaciones e incluso la detección de falacias argumentativas en las intervenciones de otros equipos. La puntuación obtenida por cada equipo se tradujo en una pequeña bonificación en la nota de la asignatura, aspecto este que ya había sido anticipado en la convocatoria de la actividad, con el fin de conseguir un mayor número de participantes en la misma.

A lo largo del desarrollo de la actividad, fueron apareciendo multitud de temas relacionados con la doble cara de las técnicas de ML e IA: sus indudables aportaciones para el avance de la humanidad en muchos ámbitos, pero también los potenciales sesgos, riesgos y amenazas de todo tipo que pueden surgir de su empleo indiscriminado o malicioso.

Finalmente, un representante de cada bando intervino en el denominado “minuto de oro” para intentar hacer una síntesis de argumentos de la posición “a favor” o “en contra”.

Tal y como se desprende de las respuestas obtenidas mediante un cuestionario pasado a los alumnos, que se analiza en la siguiente sección, los alumnos valoraron notablemente que se realizase esta actividad extra y consideraron que les ha supuesto una mejora en su nivel de conocimientos sobre los aspectos éticos y humanísticos y, sobre todo, una mejora en el grado de interés y actitud hacia los mismos.

Para los próximos cursos puede ser conveniente limitar el abanico de temas a debatir, a fin de que pueda entrarse más en profundidad en los mismos, siendo una buena idea comunicárselos con antelación a los participantes, a fin de que puedan realizar una mejor documentación previa.

También, atendiendo al volumen de participación, pueden realizarse varias rondas de debates, a fin de dar la oportunidad a un mayor número de intervenciones por parte de todos los equipos, aspecto que tuvo que estar muy limitado en esta primera experiencia, dada la enorme cantidad de participantes y el tiempo limitado de la actividad (2 horas). También podría plantearse aumentar este tiempo para poder realizar una síntesis integradora al final de la sesión.

Finalmente, puede tener sentido plantearse realizar una pequeña competición de debate con equipos de alumnos de las universidades jesuíticas españolas (Universidad de Deusto, Universidad de Loyola, ESADE, Ramón Llul...) en las cuales haya programas con asignaturas técnicas similares.

La segunda actividad planteada fue una conferencia de divulgación titulada “*el yin y el yang del deep learning*” para dar a conocer las nociones básicas acerca de qué es el ML y específicamente el deep learning (en adelante DL), y ofrecer algunos ejemplos de aplicación beneficiosa para los individuos, las empresas y la sociedad (el “yin”), así como mostrar algunas amenazas, riesgos y peligros de un mal uso de estas técnicas (el “yang”).

La conferencia fue impartida por cuatro ponentes, profesores de diversas asignaturas de ML en la Universidad Pontificia Comillas y se realizó en un estilo dinámico y atractivo, con el uso de diapositivas, videos y ejemplos en vivo de aplicaciones interesantes y sorprendentes con el fin de captar y mantener el interés y la expectación del público asistente. Al haberse realizado aún en un contexto de restricciones debidas a la Covid-19, la conferencia se realizó en formato bimodal, retransmitiéndose en streaming para aquellas personas que desearan asistir de manera remota.

La conferencia fue una de las actividades desarrollada en el contexto de la XXI Semana de la Ciencia y la Innovación de Madrid. Esta Semana es un evento de divulgación científica y participación ciudadana organizado anualmente por la Comunidad de Madrid (España), cuyo objetivo es involucrar activamente a los ciudadanos en la ciencia, la tecnología y la innovación.

La conferencia fue gratuita, de entrada libre y abierta al público en general. Los profesores pensamos que podría ser un buen complemento a la formación más técnica ofrecida en las clases regladas de la asignatura y por ese motivo la comunicamos y ofertamos de manera especial a nuestros alumnos. Con el fin de motivar la participación, decidimos ofrecer una pequeña bonificación en la calificación final a aquellos alumnos que asistieran de manera atenta a la misma.

Habida cuenta de que el campo del ML, especialmente el DL y la IA están en avance y transformación permanente y muy rápida, es necesario estar al corriente de las innovaciones que surgen mes a mes para incorporarlas al discurso, a fin de poder transmitir tanto la intuición técnica en la que se basan como reflexionar sobre sus aportes positivos y los riesgos añadidos que pueden suponer. De tal manera, será necesario realizar una puesta al día de la conferencia y sus materiales. Asimismo, merecerá la pena incorporar algún elemento adicional de participación del público con el fin de lograr un mayor dinamismo y enganche de los asistentes.

Se está considerando el ofrecer la actividad en posteriores ediciones de la Semana de la Ciencia, en otras Universidades e instituciones jesuíticas, así como, en un formato algo más simplificado, a Colegios e Institutos, dentro de las actividades de promoción de la Universidad.

### **Evidencia empírica a favor de la utilidad de nuestra metodología**

Para testar los conocimientos y actitudes sobre las implicaciones éticas del ML y la IA se desarrolló un breve cuestionario que se pasó a los alumnos que han cursado la asignatura de ML 1 durante el curso académico 2021/2022. La encuesta se envió por correo electrónico a los 180 estudiantes de la asignatura, obteniendo 58 respuestas (una tasa de respuesta del 32.2%). Hay que destacar que los alumnos estaban en pleno proceso de exámenes finales en el tiempo de pasar el cuestionario, y muchos de ellos se encontraban realizando intercambio en universidades de todo el mundo, lo que explica que el número de respuestas no sea tan alto como nos hubiese gustado.

Los alumnos de la muestra estaban en el tercer curso de sus estudios universitarios, con edades comprendidas entre los 20 y 22 años. El 56.9% cursaba el doble grado en



Administración de Empresas (ADE) y Business Analytics, mientras que el 43.1% cursaba el doble grado en Derecho y Business Analytics. El 50% eran hombres y el 50% mujeres.

El cuestionario estaba centrado en la opinión y valoración de tres dimensiones o aspectos:

1. Lo que saben. Qué grado de conocimientos creen los alumnos que tienen ellos mismos y la sociedad en su conjunto, tanto sobre los aspectos técnicos como sobre los aspectos éticos, riesgos y puntos “oscuros” del ML y la IA.
2. Lo importante que es añadir a los conocimientos técnicos las reflexiones éticas y humanísticas.
  - a. Para las personas de formación técnica
  - b. Para la sociedad en su conjunto
  - c. La importancia de que en los planes de estudios de carreras orientadas a lo técnico haya reflexión acerca de los aspectos éticos y las implicaciones sociales, económicas y éticas de los avances en ML e IA
3. La valoración de las actividades extra desarrolladas en la asignatura. Tanto a nivel de mejora de conocimientos como a nivel de mejora del interés y actitud respecto de los aspectos éticos y humanísticos
  - a. La actividad de debate
  - b. La conferencia sobre “el yin y el yang del Deep Learning”

Aunque el cuestionario usaba diferentes rangos en las escalas para preguntar sobre opiniones, importancia o valoraciones, para evitar el sesgo del método común (Podsakoff et al., 2003), posteriormente los resultados se han escalado en el rango 0-10 para mejor comparación e interpretación.

Los resultados fundamentales obtenidos pueden sintetizarse de la siguiente manera. Respecto de la dimensión de conocimientos, la tabla 1 muestra las medias y desviaciones típicas en las diferentes cuestiones planteadas.

Tabla 1. Nivel de conocimientos

		media	desviación típica
Conocimientos técnicos	alumnos	5.6	2.2
	sociedad	1.3	3.3
Formación en aspectos éticos y humanísticos	alumnos	5.0	2.9
	sociedad	1.7	2.8

Puede apreciarse que los alumnos consideran que su formación promedio en conocimientos técnicos (M=5.6) es ligeramente mejor que su formación en los aspectos éticos y humanísticos asociados a dicho campo de conocimiento (M=5.0). Asimismo, es de destacar que consideran que en la sociedad en su conjunto el nivel de formación y conocimientos tanto de tipo técnico como la relativa a los problemas éticos y puntos oscuros asociados al ML y a la IA es muy bajo y muy inferior al de los propios alumnos, alcanzando un valor promedio de 1.3 y 1.7 puntos respectivamente.

Respecto a la dimensión segunda, relativa a cómo consideran los alumnos que es importante es el conocer los riesgos sociales, económicos, ambientales y éticos en la aplicación de técnicas de ML e IA, la tabla 2 muestra los resultados resumen.

Tabla 2. Nivel de importancia de la formación es aspectos éticos y humanísticos del ML y la IA

	media	desviación típica
Para personas técnicas	7.7	2.4
Para la sociedad	6.4	2.6
Incorporación a los planes de estudio	7.8	2.3

Puede apreciarse que los alumnos dan una importancia promedio alta a que las personas técnicas tengan también formación en los aspectos éticos y humanísticos, así como en los posibles riesgos de aplicación de estas técnicas (M=7.7), considerando asimismo altamente importante que los planes de estudio de carreras orientadas a lo técnico haya suficiente formación y reflexión acerca de los aspectos éticos y las implicaciones sociales, económicas y éticas de los avances en ML e IA (M=7.8). También piensan que tiene bastante importancia que la sociedad en general tenga unos mínimos de formación en este campo (M=6.4).

Respecto a la tercera dimensión consultada, las actividades desarrolladas en el marco de la asignatura para intentar ofrecer formación y reflexión adicional en el campo ético y humanístico, la tabla 3 muestra la valoración que dieron los alumnos a estas actividades como adecuadas para conseguir dichos fines, así como su grado de participación en las mismas.

Tabla 3. Valoración de las actividades desarrolladas en la asignatura

		media	desviación típica
Debate	Valoración	6.1	2.7
	Asistencia (%)		51.7%
Conferencia Deep Learning	Valoración	6.5	2.3
	Asistencia (%)		39.7%

Los datos muestran que ambas actividades recibieron una valoración notable, pareciendo un poco más adecuada para los fines de formación y reflexión adicional en el campo ético y humanístico la actividad de la conferencia sobre DL (M=6.5) que la actividad de debate (M=6.1).

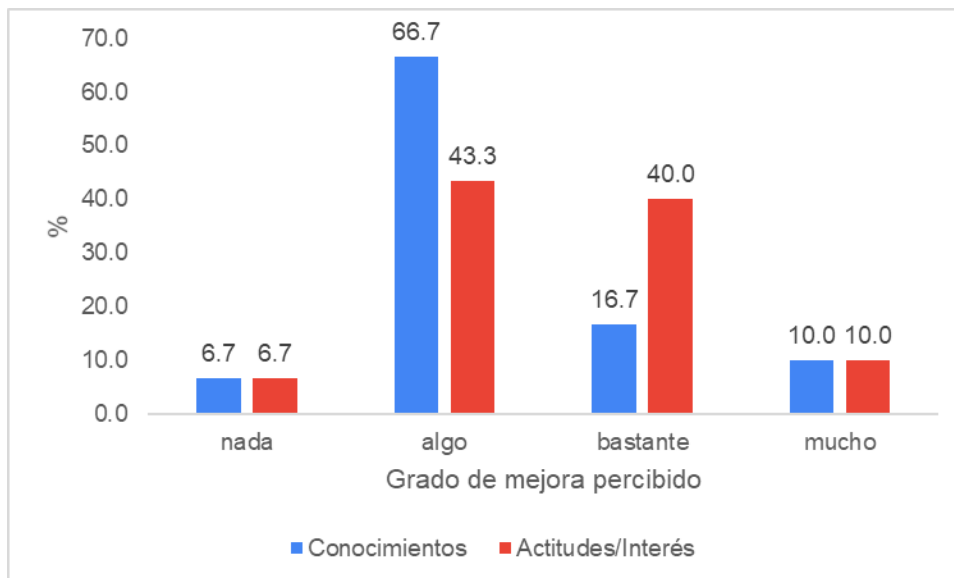
Merece la pena destacar que el debate alcanzó mucho mayor valoración media entre aquellas personas que efectivamente participaron en el mismo (M=7.0) que entre las que no (M=5.2). Esto puede indicar que es una actividad en la que es difícil estimar su grado de adecuación a priori sin participar en ella. También la conferencia, como actividad para mejorar la formación y reflexión sobre todos los aspectos éticos y humanísticos asociados a los avances técnicos en el campo del ML y la IA pareció más adecuada a los que asistieron a ella (M=6.8) que a los que no (M=6.2), aunque en este caso las diferencias de valoración son menores.

Hay que recordar que ambas actividades fueron de carácter voluntario y realizadas fuera del horario de clases habitual; a pesar de ello, el 51.7% de los alumnos de la muestra participaron en la actividad de debate y el 39.7% asistieron a la conferencia sobre DL<sup>1</sup>.

Las figuras 2 y 3 resumen el grado en el que creen que mejoraron sus conocimientos y su interés y actitud por los aspectos éticos y humanísticos aquellos alumnos que participaron en las mismas.

Respecto a la actividad de debate, se observa que el 26.7% de los alumnos participantes percibieron bastante o mucha mejora en su nivel de conocimientos, mientras que el 66.7% declaró que solo hubo algo de mejora. Sin embargo, en lo tocante al interés y la actitud por los aspectos éticos, la mitad de los participantes percibió basten o mucha mejora. Los resultados sugieren, por lo tanto, que la actividad de debate es más efectiva en la mejora de los aspectos actitudinales que en la de los cognitivos.

Figura 2. Grado de mejora en conocimientos y actitudes sobre aspectos éticos tras la celebración de la actividad de debate



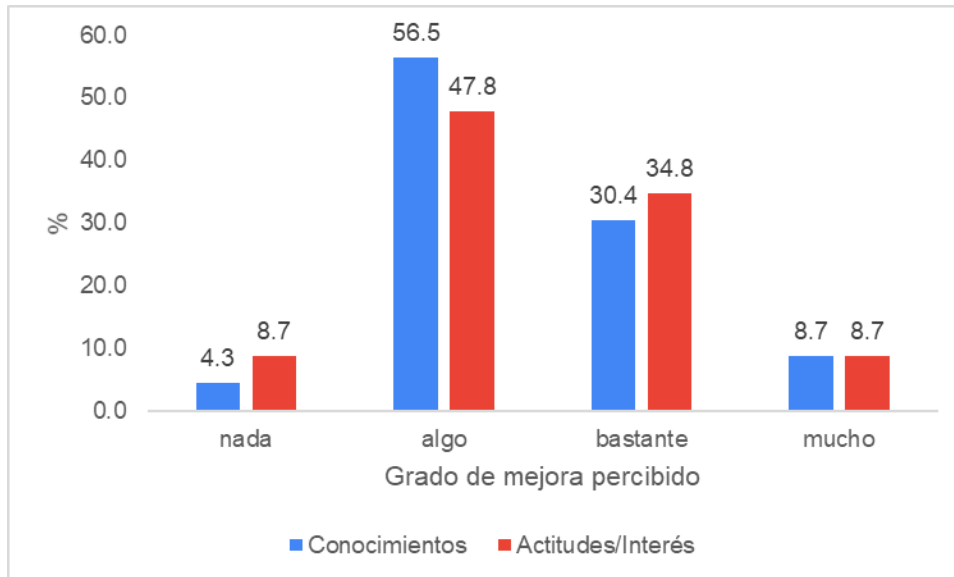
Respecto a las percepciones de mejora en los asistentes a la conferencia sobre “el yin y el yang del Deep Learning”, los datos muestran que la mejora fue pareja tanto a nivel de conocimientos como de interés y actitudes. Así, el 39.1% manifestó bastante o mucha mejora en su nivel de conocimientos, porcentaje que asciende al 43.5% en lo tocante a interés y actitud.

Los resultados sugieren por lo tanto que, la actividad de debate, al suponer una participación activa del alumno, ya desde la fase de preparación del mismo, incide más en los aspectos

<sup>1</sup> Es preciso indicar que dicha conferencia, aunque estaba abierta al público en general, se realizó en una fecha en la que los alumnos del doble grado en Derecho y Business Analytics aún no habían comenzado su docencia de Machine Learning, por lo que no se les ofreció en clase de forma directa la posibilidad de asistir, ni de obtener ninguna bonificación en la nota final de la asignatura. El porcentaje de asistentes respecto únicamente de los alumnos del doble grado en ADE y Business Analytics es del 66.7%

actitudinales que la actividad de la conferencia, en la cual hay menor participación activa del alumno. Por el contrario, la conferencia consigue mayores mejoras en los aspectos cognitivos que la actividad de debate. Es posible considerar ambas actividades como complementarias.

Figura 3. Grado de mejora en conocimientos y actitudes sobre aspectos éticos tras la celebración de la actividad de la conferencia sobre “el yin y el yang del Deep Learning”



### Conclusiones y trabajo futuro

Este trabajo constituye una memoria de la reflexión realizada por los profesores y alumnos de la asignatura de ML sobre las repercusiones éticas y filosóficas del espectacular avance de las aplicaciones de la IA en el tejido empresarial y social. En concreto, la motivación para celebrar esta reflexión ha sido la de aportar un crítico enfoque humanístico a la asignatura de ML, necesaria ante el cambio de paradigma que nuestra sociedad está experimentando ante el auge de la implementación de la IA. En particular, esta reflexión ha sido impartida por vías de un debate en la asignatura y de una actividad divulgativa expuesta en el marco de la Semana de la Ciencia 2021 a la que los alumnos pudieron asistir. Concretamente, los alumnos que han disfrutado de esta reflexión cursan en las especialidades “ADE y Business Analytics” y “ADE, Derecho y Business Analytics” de la Facultad de Ciencias Económicas y Empresariales de ICADE, centro perteneciente a la Universidad Pontificia de Comillas.

Dados los resultados de la encuesta proporcionada a los alumnos y expuestos los motivos por los que consideramos que la reflexión que hemos llevado a cabo es crítica y fundamental, para que los alumnos entiendan las repercusiones éticas y filosóficas de la práctica del ML, recomendamos la celebración de debates y actividades divulgativas en aquellas facultades en las que se impartan las asignaturas de ML, IA o DL. El propósito de esta celebración sería que el alumno no solo obtenga un nivel técnico adecuado para poder implementar estas técnicas en el ámbito empresarial sino un nivel humanístico también a la altura de poder gestionar las repercusiones éticas y filosóficas que la implementación de estas técnicas genera en la sociedad.

## Referencias

- Block, N. (1995). Some concepts of consciousness. *Sciences*, 18(2), 1-28.
- Bostrom, N. (2005). Transhumanist values. *Journal of philosophical research*,30(Supplement), 3-14.
- Candelieri, A., Perego, R., & Archetti, F. (2021). Green machine learning via augmented Gaussian processes and multi-information source optimization. *Soft Computing*, 25(19),12591-12603.
- Dixon, M. F., Halperin, I., & Bilokon, P. (2020). *Machine learning in Finance* (Vol. 1170).Springer International Publishing.
- Francisco, P. (2015). *Laudato si*. Roma: Libreria Editrice Vaticana.
- Garrido Merchán, E. C. (2021). Transhumanismo y Consciencia Fenoménica.
- Haslam, N. (2006). Dehumanization: An integrative review. *Personality and social psychologyreview*, 10(3), 252-264.
- Lai, J. P., Chang, Y. M., Chen, C. H., & Pai, P. F. (2020). A survey of machine learning modelsin renewable energy predictions. *Applied Sciences*, 10(17), 5975.
- Kulshreshth, A., Anand, A., & Lakanpal, A. (2019, October). Neuralink-an Elon Musk start-upachieve symbiosis with artificial intelligence. In *2019 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS)* (pp. 105-109). IEEE.
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *nature*, 521(7553), 436-444.
- Patterson, D., Gonzalez, J., Le, Q., Liang, C., Munguia, L. M., Rothchild, D., ... & Dean, J.(2021). Carbon emissions and large neural network training. *arXiv preprint arXiv:2104.10350*.
- Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Commonmethod biases in behavioral research: a critical review of the literature and recommended remedies. *Journal of applied psychology*, 88(5), 879.
- Ramesh, A., Dhariwal, P., Nichol, A., Chu, C., & Chen, M. (2022). Hierarchical text-conditionalimage generation with clip latents. *arXiv preprint arXiv:2204.06125*.
- Shailaja, K., Seetharamulu, B., & Jabbar, M. A. (2018, March). Machine learning inhealthcare: A review. In *2018 Second international conference on electronics, communication and aerospace technology (ICECA)* (pp. 910-914). IEEE.
- Universidad Pontificia Comillas. (2022) <https://www.comillas.edu/noticias/34-comillas-icade/icade-cee/icade-cee-ibs/476-mision-objetivos>