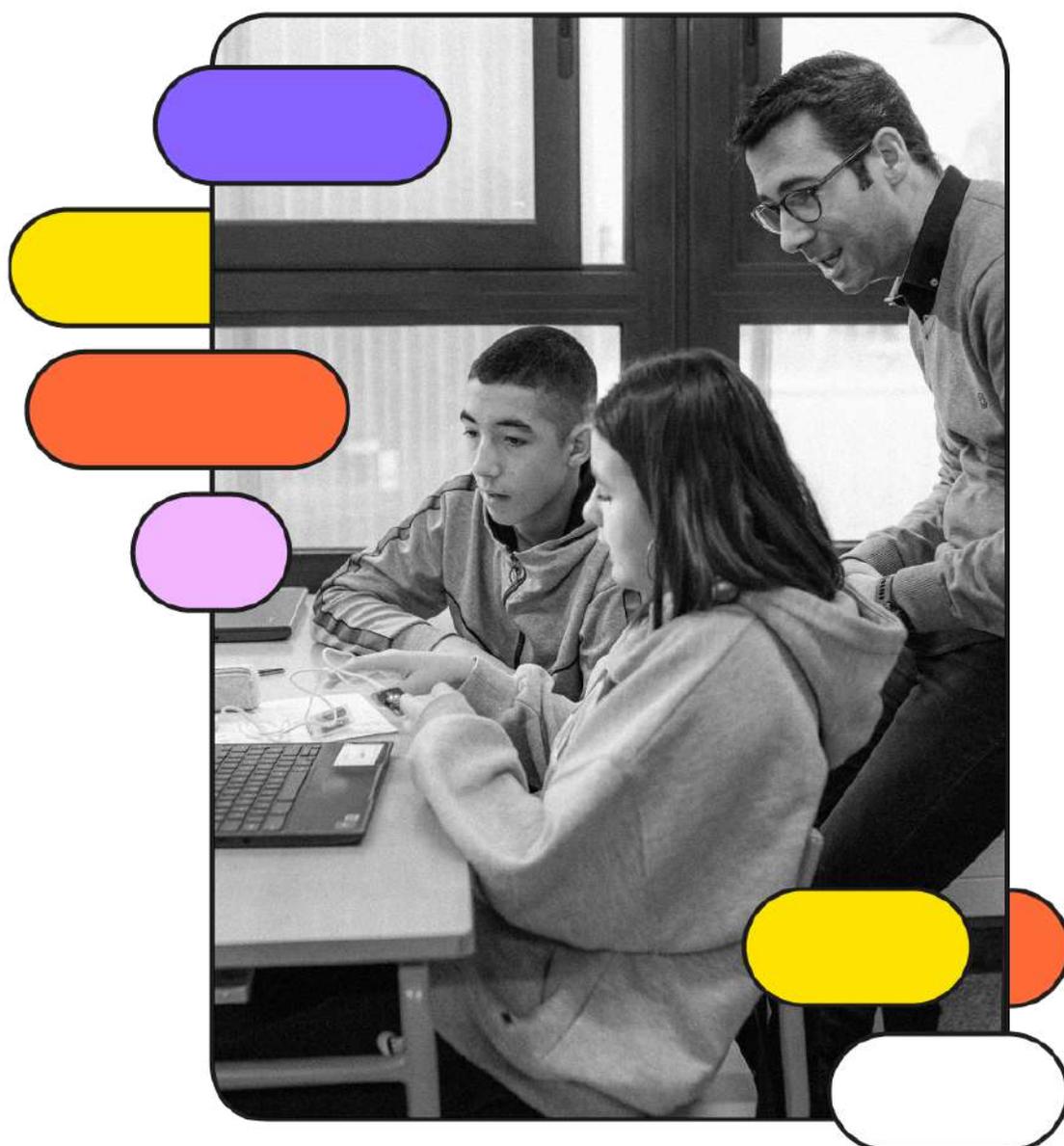


EVALUACIÓN 4.0

**Progresada adecuadamente,
pero necesita mejorar.**

Informe final del proyecto Evaluación 4.0



EVALUACIÓN 4.0

Progresada adecuadamente, pero necesita mejorar.

Informe final del proyecto Evaluación 4.0

1a Edición, marzo 2023

ISBN: 978-84-126532-1-2

© Fundació Jaume Bofill, 2022

Girona, 34

08010 Barcelona

fbofill@fbofill.cat

fundaciobofill.cat

Redacción de contenidos / Alejandra López de Aberasturi i Gómez

Coordinación de contenidos / Héctor Martínez

Dirección de contenidos / Héctor Gardó Huerta

Diseño gráfico y maquetación / Sonia Fernández

Coordinación editorial / Anna Sadurní

Edición / Fundació Bofill



Creemos que el conocimiento debe compartirse. Por ello utilizamos una licencia Creative Commons Reconocimiento 4.0 Internacional (CC BY 4.0). Os animamos a copiar, redistribuir, remezclar o transformar y crear contenidos propios a partir de esta publicación, para cualquier finalidad, incluida la comercial. Sólo os pedimos que reconozcáis la autoría de la creación original.



ÍNDICE

Introducción y contexto	4
Relación con la industria 4.0	5
Objetivos del proyecto	8
Experiencia de los y las participantes	8
Resultados derivados del análisis de datos	11
Mejoras técnicas para una mejora del rendimiento del algoritmo	16
Aprendizajes y limitaciones del consorcio impulsor	17
Recomendaciones en clave de política pública para el uso de la IA en educación	19
Conclusiones y retos pendientes	22
Nuevas líneas de trabajo y aspiraciones futuras	23
Agradecimientos	25
Anexos	26
Referencias	40

Introducción y contexto

La [masificación y diversidad de las aulas públicas y la implementación de la metodología de evaluación continua](#) han supuesto un gran desafío para la correcta conciliación de las tareas derivadas de este tipo de evaluación (por ejemplo, la corrección de ejercicios realizada por los y las estudiantes), que permiten al cuerpo docente mantener un seguimiento personalizado de su alumnado y garantizan así una educación de calidad. Entre otras muchas herramientas tecnológicas, los algoritmos de evaluación automática buscan apoyar al profesorado en esta labor. Además, estas herramientas pueden ser utilizadas en diferentes modelos educativos, tales como el presencial, el modelo en línea o, incluso, [modelos híbridos](#), aumentando su



flexibilidad y reduciendo las inequidades educativas¹. Sin embargo, **la evaluación automatizada y la retroalimentación de tareas de respuesta abierta siguen siendo un desafío tecnológico por lo complejo de automatizar los procesos de evaluación**, a pesar de los avances recientes en el campo del procesamiento del lenguaje natural². **Y estos procesos de evaluación son de vital importancia, dado que son clave en todos los procesos de aprendizaje.** Además, es importante tener en cuenta **las limitaciones de las herramientas de inteligencia artificial en tareas creativas** o artísticas, lo que nos lleva a cuestionarnos su capacidad para llevar a cabo procesos de evaluación ricos, diversos y fiables. Cuando hablamos de un proceso de crecimiento educativo, **consideramos que la evaluación es aquel momento en el que aprendemos más y mejor**, por lo que tenemos que poner **en el centro del debate la posibilidad de que los algoritmos puedan, o no, ayudar a la comunidad educativa en dicha tarea, y hacerlo de manera imparcial, objetiva, honesta y sin prejuicios ni sesgos de ningún tipo.**

¹ En esta línea, recomendamos consultar el artículo "[La escuela híbrida, más allá de ordenadores e internet](#)", derivado del informe [Com impulsar la transformació digital de l'escola](#), que analiza el potencial impacto de la hibridación de la educación, en clave de equidad, y las posibilidades de generar oportunidades significativas y de calidad para los y las estudiantes que más lo necesiten en el contexto educativo de Catalunya.

²Los algoritmos de procesamiento del lenguaje natural (PLN) son aquellos que se encargan de entender, procesar e intentar generar lenguaje tal y como lo haría una persona. Un ejemplo claro de esto ello los *chatbots*, actualmente (enero de 2023) tan en auge, como podría ser Chat GPT.

Además de los desafíos técnicos y teóricos, [existe una dimensión ética en el uso de estas herramientas](#), ya que la evaluación tiene un impacto en el desempeño y el futuro del alumnado. Aunque se ha avanzado en hacer las redes neuronales más transparentes, aún existen limitaciones en la explicabilidad³ de las decisiones tomadas por estos algoritmos. Por esta razón, **muchos investigadores/as proponen soluciones híbridas que combinan la automatización con la intervención humana**, como por ejemplo la evaluación entre pares. Además, estudios en psicología indican que [la evaluación entre pares tiene un impacto positivo en el proceso formativo de los y las estudiantes](#), ya que fomenta la autorreflexión. En este sentido, se estudian y comparan dos soluciones híbridas que utilizan un enfoque probabilístico para estimar la distribución de probabilidad de cada evaluación automática. Estas entradas son las evaluaciones entre compañeros/as de una clase (coevaluación) y un pequeño porcentaje de las calificaciones del o la docente (lo que se conoce como *ground truth*⁴).

Relación con la industria 4.0

En el contexto de la educación, la industria 4.0 ofrece un gran potencial para mejorar la tecnología desplegada en el aula. Sin embargo, para que esta mejora sea posible, **es esencial que la industria y los agentes de I+D+i tengan acceso a datos precisos y relevantes como los obtenidos en la presente investigación**. También pensamos que se debe promover el estudio y el **desarrollo de herramientas de inteligencia artificial (IA) que potencien las capacidades humanas sin agregar sesgos ni perjudicar a quienes presenten condiciones iniciales vulnerables**. Concretamente, urge evitar el uso de modelos preentrenados con datos masivos para hacer predicciones sobre el desempeño que tendrán los y las estudiantes en sus exámenes o [la idoneidad de un](#)

³ Definición obtenida del sistema OpenAI sobre el término *explicabilidad*: “La explicabilidad en inteligencia artificial se refiere a la capacidad de entender y dar una explicación clara y comprensible sobre cómo y por qué un sistema de IA toma una determinada decisión o realiza una determinada acción. Esto es importante porque muchos sistemas de IA utilizan métodos complejos y opacos para tomar decisiones, lo que puede generar incertidumbre, desconfianza y, en algunos casos, discriminación injusta. La explicabilidad es un aspecto clave de la ética en IA y es esencial para asegurar que los sistemas de IA sean responsables, transparentes y fiables”.

⁴ Definición obtenida del sistema OpenAI sobre el término *ground truth*: “*Ground truth* es un término utilizado en el ámbito de la inteligencia artificial y el aprendizaje automático. Se refiere a la verdad real o la información confiable y verificable sobre un determinado conjunto de datos. Es decir, es la información que se utiliza como referencia para evaluar la precisión de un modelo de aprendizaje automático o para etiquetar datos. Esta información es esencial para entrenar y evaluar modelos de aprendizaje automático y para garantizar que estos modelos produzcan resultados precisos y confiables”.

Vamos a poner un ejemplo para intentar hacerlo más comprensible: en una base de datos de mamografías usada para entrenamiento de un algoritmo de detección precoz de cáncer, cada imagen iría acompañada de una etiqueta (*i. e.* cáncer/no cáncer o positivo/negativo). A esta combinación “imagen+etiqueta” es a lo que llamamos *ground truth* o regla de oro. De este modo, la red neuronal va siendo capaz de distinguir los rasgos comunes a todas las imágenes que se corresponden con la etiqueta “cáncer”.

[candidato/a que aplica a una universidad](#), ya que esto puede llevarnos completamente a lo opuesto, a generar más inequidad o a condicionar las decisiones de estos (lo que se llama [el efecto Pigmalión](#)). Es decir, las expectativas que nuestro entorno deposita en nosotros pueden condicionar nuestra toma de decisiones, lo que, en algunos casos, podría no ser beneficioso (si un algoritmo nos propusiera qué estudiar, según los datos de nuestra carrera estudiantil, ¿deberíamos hacerle caso o, por el contrario, deberíamos escoger el itinerario formativo que más deseamos?). Además, **es importante que la Administración esté informada y tenga la capacidad de informar, y formar, a la comunidad docente para que pueda solicitar herramientas tecnológicas que no vulneren los derechos de sus estudiantes** (ver más abajo el apartado de [políticas públicas](#)). Sería deseable, pues, que los tomadores/as de decisiones —Administración, docentes, desarrolladores/as, pedagogos/as y psicólogos/as— **trabajasen juntos para crear herramientas que se adapten a las necesidades reales del aula** y que aporten más valor al ámbito educativo.

La escuela juega un papel crucial en la formación de la ciudadanía del mañana en un mundo cada vez más tecnológico. Por lo tanto, es esencial que **el sistema educativo no sea agnóstico, tecnológicamente hablando, sino que enseñe al alumnado a utilizar las herramientas necesarias de manera efectiva y crítica, y a comprender su funcionamiento en profundidad**. Además, **es importante formar a personas conscientes de los peligros asociados a las tecnologías digitales** y dotarlas de las herramientas y la autonomía necesarias para evitar dichos riesgos, entre ellos:

- La perpetuación, en las decisiones de los algoritmos, de los sesgos que la sociedad ya presenta, asociados al fin y al cabo a las percepciones humanas que hay detrás de ellos.
- La posibilidad de generar profecías que se acaben cumpliendo, propiciando así el resultado opuesto al esperado.
- La amenaza que el empleo de la tecnología puede suponer para la privacidad.
- La pérdida de concentración que puede acarrear el uso constante de las redes sociales.
- El rastreo de la actividad y el uso de nuestros datos que realizan las empresas anunciantes.
- La polarización política en la que puede desembocar el uso exclusivo de sistemas recomendadores de contenidos en función de los intereses de las personas usuarias.
- La dismorfia corporal derivada del uso de filtros en redes sociales.

En cuanto a la tecnología, también es importante saber definir los límites de su utilización en el aula, ya que **quizás no todo deba ser digitalizado**. En algunos casos, la educación presencial y sin pantallas puede ser la mejor opción para atender a las necesidades educativas y sus particularidades, que se cubrirían mejor en contextos no virtualizados o mediados por la tecnología. Resumiendo, también hay que tener criterios específicos, pactados con todos los actores de la comunidad educativa

(Administración, equipos directivos, docentes, estudiantes y familias), que establezcan qué procesos deberían ser potencialmente objeto de regularización o automatización, en beneficio de los itinerarios educativos de los y las estudiantes, y cuáles no. El objetivo debería ser evitar la virtualización innecesaria o forzada de aquello que puede seguir funcionando desde lo analógico, manteniendo la empatía y la cercanía entre docente y alumnado, y aplicar algoritmos y tecnologías solo en aquellos procesos susceptibles de ser optimizados con su uso.

Por otra parte, no puede obviarse el hecho de que la tecnología puede tener un impacto muy positivo en la educación al acercar el conocimiento a personas usuarias que, de otra manera, no habrían tenido acceso al sistema educativo (por ejemplo, mediante algunas experiencias y técnicas basadas en la [educación híbrida](#)).

En definitiva, **la escuela de hoy en día, contemporánea, debe equipar a los futuros actores de nuestra sociedad** con criterios para que sepan utilizar, comprender y reivindicar tecnologías inteligentes, y también protegerse de ellas. Con respecto al uso de estas tecnologías en la escuela, las empresas y los desarrolladores/as deben trabajar codo con codo con equipos diversos y multidisciplinares (por ejemplo, psicólogos/as, educadores/as) que guíen su buena praxis, a fin de evitar en la medida de lo posible los riesgos psicológicos asociados al uso de tecnologías inteligentes y aprovechar las ventajas que ofrecen.



Objetivos del proyecto

[Evaluación 4.0](#) ha testado una solución de inteligencia artificial (IA) aplicada a los procesos de evaluación educativa **para explorar cuáles son los riesgos y las oportunidades de la IA en contextos educativos**.

El piloto se llevó a cabo mediante la prueba de un algoritmo que puede permitir reducir el tiempo dedicado por la comunidad docente a procesos de evaluación. El algoritmo pretende agilizar la labor del profesorado, buscando las similitudes entre las evaluaciones llevadas a cabo por el alumnado y por el cuerpo docente. En caso de conseguirse, esta reducción de los tiempos en tareas de menor valor añadido durante procesos evaluativos [podría ayudar a los y las docentes a reducir el tiempo dedicado a revisar ejercicios, permitiéndoles más tiempo de calidad en el acompañamiento personalizado](#) del alumnado.

El proyecto se impulsó con la voluntad de experimentar y cuestionar la implantación de la IA en la educación, y estudiar sus potenciales ventajas para mejorar las formas de aprender.

En el aspecto técnico, Evaluación 4.0 quiere validar que las evaluaciones ejecutadas por PAAS (*personalised automated assessments*, o evaluaciones personalizadas automatizadas) [1] tengan una “calidad” equivalente a las evaluaciones realizadas por el profesorado en proyectos colaborativos. **Se busca medir la calidad de estos algoritmos en términos de cercanía a las evaluaciones del o la docente**, así como en la reducción del tiempo de evaluación. Esto es, se busca comprobar si las evaluaciones calculadas por los algoritmos serían equivalentes a las realizadas por el o la docente, y si además reducirían el tiempo de corrección. Sin embargo, es importante tener en cuenta que **estas soluciones híbridas deben ser cuidadosamente diseñadas y monitoreadas para [asegurar que se cumplan los estándares éticos y educativos apropiados](#)**, y que los y las estudiantes reciban retroalimentación útil y precisa.

En resumen, **las herramientas de evaluación automática pueden ser una gran ayuda para el profesorado, pero es crucial considerar las limitaciones y los desafíos éticos y educativos asociados con su uso**. Además, es fundamental tener en cuenta soluciones híbridas que combinen los beneficios de la automatización con la intervención humana y la retroalimentación para garantizar un proceso educativo eficaz y equitativo.

Experiencia de los y las participantes

Una vez finalizado el piloto, se preguntó a los centros participantes acerca de su visión sobre las nuevas oportunidades que puede ofrecerles la tecnología, y en particular la inteligencia artificial, en sus proyectos educativos. Se invitó a participar en esta valoración a los y las docentes líderes de los diferentes grupos y, al mismo

tiempo, se les invitó a compartir dicho formulario con las personas responsables de los centros en los que ejercen (entre otros, perfiles tales como directores/as de centro y/o coordinadores/as pedagógicos/as). Empezamos pidiéndoles que nos citaran **un posible temor ante el uso de la IA en la educación** y, en particular, en sus centros educativos. Sobre este temor, nos mencionaron que:

- En el proceso de utilizar algoritmos para gestionar los datos del alumnado, **consideran importante que las empresas responsables cumplan con las normativas de privacidad y se establezca un entorno seguro para su implementación**, especialmente en el caso de estudiantes con necesidades educativas especiales.
- **Los algoritmos deberían tener en cuenta las distintas variables obtenidas para evitar desigualdades** y se debería hablar claramente sobre ello a la comunidad educativa (estudiantes, familias y docentes) para evitar generar más desconfianza.
- Nunca se debe perder la empatía hacia el alumnado ni **el trato humano, y siempre hay que mantener la empatía en el ámbito educativo**.

En general, creen que **el mayor problema no es discutir si la IA se debe implementar o no en la educación, puesto que la mayoría lo considera ya una realidad, sino** que el dilema moral está en estudiar y debatir acerca de los posibles usos indebidos, y ocultos, de la IA en contextos educativos.

Les pedimos que nos mencionaran cuáles creían que eran las potencialidades de usar IA en educación para intentar ofrecer mejores oportunidades educativas. Ante esta cuestión:

- Ven en su implementación **la posibilidad de ofrecer un aprendizaje más personalizado, que permitiría mejorar la eficiencia del trabajo docente**, ya que el profesorado tendría que invertir menos tiempo en cuestiones como la creación de grupos o la evaluación de ciertas tareas y podría dedicarse a la personalización del aprendizaje de su alumnado.
- Valoran el crecimiento y desarrollo de proyectos colaborativos, puesto que consideran que **ofrece diferentes actividades de enseñanza-aprendizaje motivadoras para los y las estudiantes, de modo que pueden brindarles el protagonismo que necesitan** para que el aprendizaje se vuelva vivencial y significativo.
- Mencionan **la importancia de la personalización del aprendizaje como una clara mejora de la atención a la diversidad**, que les permite adaptar su trabajo a las necesidades y particularidades de cada alumno/a.
- Piensan que sería posible **utilizar ciertos indicadores predictivos para apoyar su misión educadora y de acompañamiento al alumnado**, reforzando procesos de tutorías y apoyo a la creación de itinerarios formativos.
- Ven en la IA una **ayuda relevante para conseguir evaluar de manera más eficiente**, seguramente reduciendo tiempos de dedicación a procesos que, en

alguna ocasión, pueden ser operativos, y teniendo más tiempo para procesos que requieren de su empatía, cercanía y parte más humana para conseguir resultados relevantes (tutorías, acompañamiento, personalización del aprendizaje...).

- Consideran que **podría ser una ayuda muy importante poder disponer en tiempo real de información sobre cómo está aprendiendo y comportándose su alumnado**: qué temáticas sería mejor reforzar, en qué competencias habría que hacer más hincapié y con qué grupo específico de estudiantes, o cuánto tiempo le lleva a cada estudiante alcanzar los objetivos de aprendizaje propuestos, entre mucha otra información que haría de su tarea un proceso más diverso y enriquecedor.

Quisimos hacer una prospección sobre cómo ven la situación actual de sus centros educativos para **conocer cuáles son los retos más complejos e importantes a los que se enfrentan a la hora de realizar cambios sistémicos en su forma de trabajar**. Ante esta cuestión, nos hicieron saber que:

- Aunque creen interesante la aplicación de la IA para facilitar la personalización del aprendizaje y la evaluación automática de procesos, **hay una manifiesta preocupación por la privacidad de los datos de los y las estudiantes y sus familias**.
- **Ven complejo que se pueda aplicar a todas las materias**, ya que se tendrían que adaptar las actividades, lo que implicaría vencer muchas reticencias y metodologías. Un posible experimento podría ser que el profesorado interesado pudiera evaluar un trimestre con algoritmos y otro sin IA para hacer una comparativa, e intentar conocer en profundidad su aplicación en el día a día.
- Opinan que **es necesario un mínimo de formación docente para usar estas herramientas y generar confianza en su utilización** para así poder transmitirla al resto de sus respectivos claustros.

Finalmente, **les preguntamos cuáles opinan que son las cuestiones éticas y morales vinculadas al uso de la IA en entornos educativos que cabría debatir previamente a su implementación** en procesos de aprendizaje. En esta cuestión:

- Les preocupa, **como aspecto primordial a tener en cuenta, la privacidad y seguridad de los datos más sensibles**, especialmente los provenientes de colectivos estigmatizados, así como abordar el uso que se hace de los datos de este alumnado.
- **Consideran que la figura del docente es insustituible**, puesto que el acompañamiento y la empatía no se pueden programar. En este sentido, hacen hincapié en la posibilidad de usar la inteligencia artificial como recurso para **una evaluación más personalizada, siempre y cuando se garantice una corrección justa y se tengan en cuenta los peligros que implica la recopilación de datos privados por parte de la IA**.

- Finalizan sus reflexiones **haciéndonos saber que es importante que el alumnado entienda que la evaluación mediante IA es igual de justa que la de cualquier método convencional**. La personalización de los comentarios y las observaciones puede ser una ventaja añadida que les ayude en sus itinerarios formativos, complementando la evaluación hecha por la IA.

Resultados derivados del análisis de datos

El objetivo de este estudio ha sido aplicar y validar el modelo PAAS [1] para la corrección automática de exámenes hechos en clases presenciales. Para ello, utilizamos los datos de coevaluación (o evaluación entre pares, es decir, entre alumnos/as) y de heteroevaluación (o la evaluación realizada por el profesor/a a sus alumnos/as) recogidos en clases presenciales de varios centros educativos, y estimamos el error de predicción del algoritmo como función del porcentaje de evaluaciones del profesor/a introducidas en el modelo. Al igual que en estudios previos, esperábamos observar un descenso del error de predicción conforme aumentara el número de evaluaciones del o la docente conocidas por el algoritmo. Para seleccionar estas evaluaciones del profesor/a, que se introducían paulatinamente, se siguieron dos métodos: selección aleatoria y un heurístico de minimización de entropía⁵.

A este respecto, esperábamos observar una ventaja relativa del uso del heurístico con respecto a la selección aleatoria. Es decir: se esperaba que seleccionar las evaluaciones docentes (*ground truths*, GT) introducidas en el modelo siguiendo el criterio de reducción de la incertidumbre repercutiera en predicciones del algoritmo de mayor calidad que si las GT se seleccionaran aleatoriamente.

⁵ Muy brevemente, un heurístico de minimización de entropía es un criterio de selección (en este caso, de las evaluaciones del o la docente que se introducen en el sistema) que busca minimizar la incertidumbre que el algoritmo tiene sobre sus predicciones. Dado que el modelo PAAS usa coevaluaciones de alumnos/as para el cálculo de las notas automáticas, puede ocurrir que las coevaluaciones de algún ejercicio sean muy variables o que PAAS no haya aprendido aún cuánto debe confiar en el criterio de los coevaluadores/as a quienes les tocó corregir ese ejercicio. El heurístico de minimización de entropía entonces decide descubrir la nota que el o la docente puso a ese ejercicio para evitar este alto grado de incertidumbre en la evaluación.



Nuestro modelo estima una distribución de probabilidad para cada calificación automática. Para ello, recibe como entradas las coevaluaciones de compañeros/as de clase, así como un porcentaje dado de las calificaciones del o la docente (en adelante, *ground truths*, GT). El modelo PAAS [1] aplica la teoría de los sistemas multiagente, de manera que, dada una comunidad de agentes (estudiantes) y un líder (docente), cuyo criterio de evaluación se considera como el verdadero o correcto, PAAS construye una matriz que representa la confianza que el o la docente debería depositar en las coevaluaciones de su alumnado.

Esta confianza es utilizada como medida para “pesar” las coevaluaciones de los miembros de la comunidad (estudiantes), de manera que quienes presentan un estilo evaluativo más similar al del o la docente recibirán más crédito. En el caso de que no existan exámenes evaluados en común por el o la docente y un alumno/a determinado/a, PAAS hará uso de la noción de transitividad de la confianza: si el o la docente confía en el o la estudiante A en tanto que buen evaluador/a, y a su vez el o la estudiante A confía en el o la estudiante B, entonces es razonable deducir que el o la docente confiaría también en tener evaluaciones en común con B.

Sintetizando, el algoritmo intentará valorar las diferentes evaluaciones realizadas por los o las diferentes estudiantes, otorgando un mayor peso a aquellas evaluaciones que, según el criterio mencionado, más se parezcan al estilo evaluador del o la docente. De este modo, se cree posible intentar reducir el tiempo estimado de corrección por parte del o la docente, tanto en cuanto la herramienta informática podría evaluar multitud de tests solo con los datos obtenidos a partir de unas cuantas coevaluaciones entre estudiantes y heteroevaluaciones efectuadas por el o la docente.

En el caso de los experimentos llevados a cabo en este proyecto, se formaron grupos pequeños o parejas de estudiantes, de manera que cada ejercicio a evaluar era resultado de un esfuerzo colaborativo y, por lo tanto, cada docente tuvo que evaluar tantos ejercicios como grupos tuviera a su cargo. En total, 23 clases (cada una a cargo de un o una docente) participaron en el estudio. Aunque la gran mayoría de los y las docentes corrigieron la totalidad de los ejercicios de los grupos que estaban a su cargo, hay algunos casos en los que faltan datos. Desconocemos si el motivo fue un fallo de la plataforma o un fallo humano. A consecuencia de esta falta de datos, el promedio de ejercicios revisados por los profesores/as no fue el prescrito (100%), sino aproximadamente el 70%.

La media de ejercicios revisados por docente fue de entre 2 y 12 ($M= 6,4$, $SD=3,2$)⁶. Por su parte, cada alumno/a (individualmente) comprobó entre 1 y 5 ejercicios de grupos de compañeros/as ($M=1,77$, $SD=1,6$).

Una vez recogidos los datos y corregidos en su integridad, procedimos a testar el algoritmo. Para ello, introdujimos la totalidad de las coevaluaciones entre compañeros/as en el sistema (un total de 734 coevaluaciones de 118 ejercicios grupales).

Para cada clase, se siguió el siguiente procedimiento:

- Se identificaron las evaluaciones del o la docente.
- Se introdujeron paulatinamente en el modelo el 10, 20, 30, 40, 50, 60, 70, 80 y 90 % de las evaluaciones del o la docente a cargo de la clase. Para cada porcentaje p de GT, se calculó la predicción del algoritmo del $(100-p)\%$ de los ejercicios restantes. Se siguieron dos criterios diferentes para la elección de las p correcciones del o la docente: selección aleatoria (línea violeta en las figuras) y selección del ejercicio grupal que generaba máxima incertidumbre en el cálculo de la nota automática (línea azul).
- Las notas predichas por el algoritmo se compararon con las notas del o la docente. En caso de que esa nota no estuviera disponible en la base de datos, ese punto no se tenía en cuenta en el posterior análisis de error. El error se calculó como error relativo medio.

En general, las curvas de error obtenidas ([figura 1](#)) siguen un comportamiento similar al reportado en otros estudios con PAAS [\[1\]](#). A medida que el número de evaluaciones del profesor/a introducidas en el sistema aumenta, el error de predicción de PAAS

⁶ M: media de la muestra.
SD: desviación estándar de la muestra.

desciende. Tal y como esperábamos, en general este valor de error baja monótonamente, no siendo así cuando el número de evaluaciones por alumno/a es muy bajo o cuando hay gran escasez de evaluaciones del o la docente. En tales casos, se observa una tendencia inestable de la curva de aprendizaje, de manera que en ocasiones la adición de nuevas evaluaciones del o la docente induce predicciones con un error mayor que el obtenido para porcentajes de GT menores. Este resultado es esperable, dado que cuando el número de evaluaciones por alumno/a es bajo, el algoritmo tiene pocos datos para calcular correctamente el peso que debe otorgar a cada evaluador/a. Notamos un descenso del error más acelerado para clases en las que el número de evaluaciones del profesor/a disponibles es mayor. De nuevo, si pensamos en términos porcentuales, esto tiene sentido, ya que el número de GT que se introducen en cada ciclo en clases grandes es mayor. De hecho, para casos como el de la [figura 1 \(dcha.\)](#), en el que el número de *ground truths* es menor que 10, hay algunos porcentajes que se corresponden con el mismo número de evaluaciones observadas por el algoritmo. En el caso que nos ocupa, esta situación se da para los porcentajes de 30 % y 40 % y para los de 60 % y 70 %. Las diferencias observadas se deben a la semilla aleatoria con la que se inicia cada bucle.

En este sentido, los elevados valores de error cuando $p=10\%$ se entienden cuando se observa que el número de evaluaciones del o la docente disponibles nunca es mayor de 15, de modo que introducir el 10 % de ellas a menudo supone tener un único *ground truth* en el sistema.

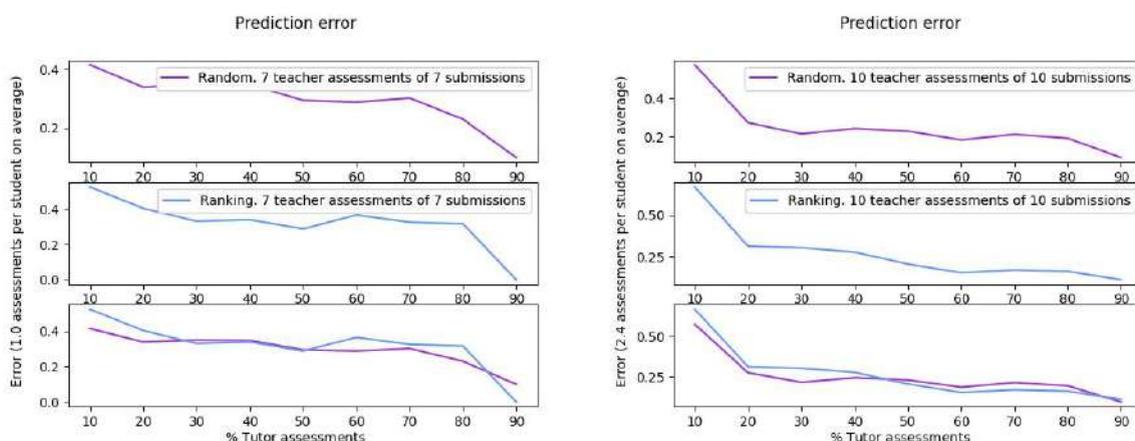
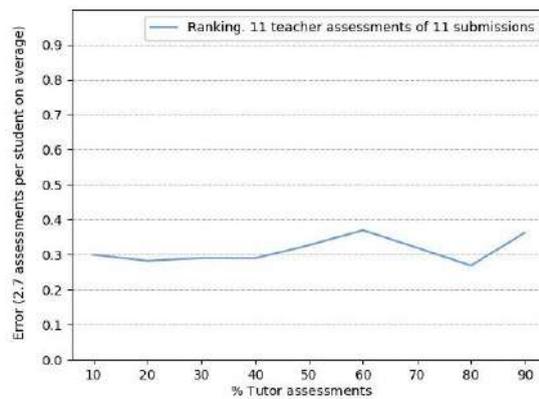
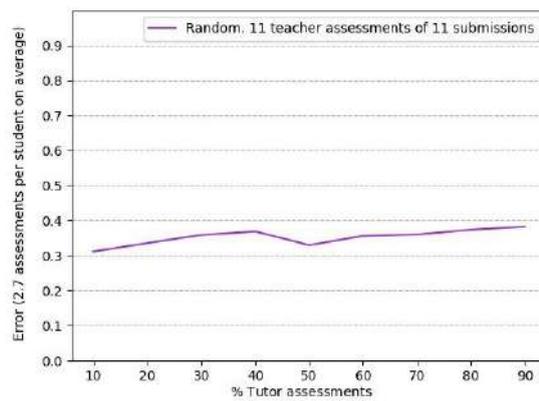


Figura 1. Evolución del error relativo medio para dos clases diferentes conforme aumenta el número de evaluaciones del o la docente introducidas en el modelo. Las líneas azules se corresponden con un criterio de selección de GT que minimiza la entropía. Las líneas violeta se corresponden con un criterio de selección aleatorio. **Izda.:** clase en la que el o la docente evaluó 7 de los 7 grupos a su cargo. El número medio de ejercicios corregidos por alumno/a fue de 1,0. **Dcha.:** clase en la que el o la docente evaluó 10 de los 10 grupos a su cargo. El número medio de ejercicios comprobados por alumno/a fue de 2,4.

Una excepción en la tendencia de las curvas descrita es el caso de una clase en la que el promedio de coevaluaciones por ejercicio fue de 2,7 y hubo 11 evaluaciones disponibles del profesor/a, aunque la curva de aprendizaje nunca llegó a descender ([figura 2](#)), lo que puso de manifiesto que el algoritmo no fue capaz de atribuir el peso correcto a los coevaluadores/as. Al inspeccionar los datos de esa clase, se observa

que se trata precisamente del grupo que presentó la máxima desviación estándar media (1,8 puntos) en la diferencia con las evaluaciones del profesor/a. Esto es: los datos de esta clase fueron tales que las diferencias entre alumnos/as y profesor/a no se mantuvieron de manera consistente: así, alumnos/as que en algunos casos daban varios puntos de más a sus compañeros/as con respecto al o la docente en otras ocasiones realizaban evaluaciones muy por debajo de las del profesor/a. Cada vez que el algoritmo añadía una nueva evaluación del o la docente, las matrices de confianza eran recalculadas y cambiaban en gran medida con respecto a la iteración anterior, de manera que el algoritmo nunca llegó a aprender los pesos adecuados para cada alumno/a.



Prediction error

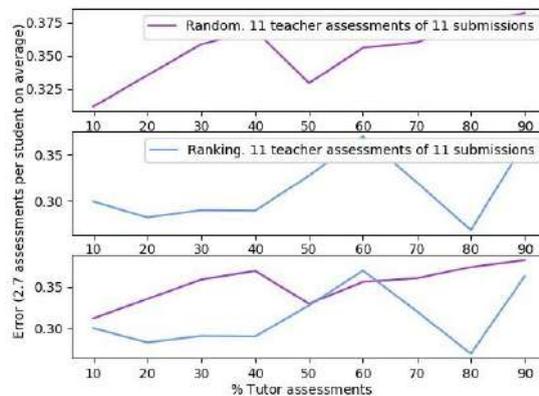


Figura 2. Curva de error relativo medio conforme aumenta el número de evaluaciones del o la docente introducidas en el modelo para una clase en la que el número de evaluaciones disponibles del o la docente fue de 11 y cada alumno corrigió 2,7 ejercicios de media. Esta clase se caracterizó por ser la que presentó la máxima desviación estándar media en las diferencias entre alumnos/as y docente en la evaluación, por lo que el algoritmo no fue capaz de asignar correctamente los pesos a los diferentes coevaluadores/as y el error nunca bajó.

De hecho, observamos que cuanto más baja es la desviación estándar media en las diferencias de evaluación entre alumnos/as y docente, tanto más siguen las curvas de error la forma original reportada en [1]. Nuevamente, este resultado es lógico, puesto que PAAS será más preciso para aquellas clases en las que más se manifieste el principio de confianza transitiva y en las que el patrón de diferencias de corrección entre alumnos/as y docente se mantenga más estable.

De manera similar a estudios anteriores, en algunos casos observamos una leve ventaja del heurístico de minimización de entropía (línea azul) sobre la selección aleatoria de evaluaciones del o la docente. Cuando esa ventaja se da, se produce un cruce de las líneas y la azul queda por debajo (menor error). Este punto de corte varía en cada caso, y no queda claro cuáles son las variables (tamaño de la clase, número de evaluaciones por estudiante, número de evaluaciones del o la docente disponibles) que lo determinan.

Por otra parte, y sin perder de vista las dificultades que ello supone, creemos que un número mayor de evaluaciones por clase de los o las docentes nos habría permitido realizar una comparativa de los resultados con estudios previos más completa, y probablemente las curvas de error habrían tenido un comportamiento más similar al canónico. En este sentido, sería deseable para el futuro que el número de grupos de trabajo por clase fuera mayor (o que se trabajase de manera individual), ya que cuantas más muestras hay en el sistema sobre cómo evalúa el o la docente, mayor es el incremento de *ground truths* que se introducen por iteración y más informativo es el correspondiente cambio en el error de las predicciones (véase con más detalle el apartado [Conclusiones y retos pendientes](#), donde se sintetizan los resultados obtenidos del estudio y sus conclusiones).

Mejoras técnicas para una mejora del rendimiento del algoritmo

A la luz de los resultados anteriores, proponemos las siguientes vías de actuación para la mejora del algoritmo:

1. Explorar *a priori* otras distribuciones de probabilidad para representar las diferencias entre evaluadores/as que aceleren el descenso de la curva de error: actualmente, cuando dos evaluadores/as no tienen ninguna evaluación en común, suponemos que las diferencias de evaluación entre ellos o ellas se distribuirán uniformemente. Es decir, suponemos que si tomamos a dos personas al azar que van a comprobar el mismo ejercicio, de entrada, no hay

motivos para creer que la diferencia entre sus notas tendrá un valor preferente: hay la misma probabilidad de que se diferencien en el máximo de puntos posibles (uno puntúa con el mínimo y otro puntúa con el máximo) que de que no se diferencien en absoluto (ambos coinciden en la nota). En el futuro, sería interesante explorar el rendimiento del algoritmo si sustituimos esta hipótesis de diferencias uniformes por otra. Por ejemplo, podríamos verificar cómo cambia la calidad de las predicciones si partimos de la idea de que es más probable que dos evaluadores/as al azar coincidan en la nota con un cierto margen de variabilidad que de que pongan notas diametralmente opuestas (p. ej., diferencias entre evaluadores/as distribuidas como [una campana de Gauss](#)).

2. Explorar otros heurísticos (además de la minimización de la entropía) para la selección automática de GT en cada ciclo de evaluación: hasta ahora habíamos comparado el rendimiento de PAAS cuando se introducían *ground truths* al azar vs. cuando se introducían para intentar reducir en cada iteración la incertidumbre del algoritmo. En el futuro, podríamos utilizar otros heurísticos (reglas) de selección (p. ej., selección del GT correspondiente al ejercicio que fue coevaluado por un mayor porcentaje de alumnos/as a los que el algoritmo da poco crédito).
3. Estudiar formas alternativas de asignar ejercicios a evaluadores/as que maximicen la difusión de confianza transitiva en la red de agentes, evitando así que haya clústeres de agentes aislados o que tienen pocos nexos de confianza con el resto: como se ha explicado, PAAS calcula *quién se parece a quién* corrigiendo. Concretamente, si un alumno/a se parece al o a la docente, y otro alumno/a se parece a ese alumno/a, entonces PAAS trabaja bajo la premisa de que es altamente probable que ese alumno/a también se parezca al profesor/a, aunque no haya pruebas directas de ello. Sin embargo, ¿qué pasa cuando tenemos un alumno/a B que no tiene evaluaciones en común con el o la docente y además tiene pocas evaluaciones en común con otros compañeros/as? Decimos que B está aislado en la red de confianza (hay pocos nexos que lo unen a los demás). En ese caso, calcular el peso que debería dársele a la opinión de B resulta una tarea altamente incierta, puesto que hay poca muestra con la que calcular su perfil como evaluador/a. Para evitar este problema, sería de gran interés optimizar la elección de *quién evalúa qué* para que la red de confianza sea lo más *tupida* posible y la confianza transitiva ([A confía en C] porque [A confía en B] y [B confía en C]) pueda fluir, sin dejar aislado a ningún alumno/a.

Aprendizajes y limitaciones del consorcio impulsor

Este ha sido un proyecto que **ha puesto de manifiesto las dificultades existentes a la hora de acercar tecnologías de vanguardia a la comunidad educativa**, dados los distintos ritmos de trabajo y estilos organizacionales. Tal y como anticipamos en el diseño del proyecto, creemos que **uno de los mayores mecanismos para transformar**

la sociedad es la educación, y consideramos la tecnología como un elemento potenciador que la puede ayudar a ser más eficiente, eficaz y motivadora. Asimismo, somos conscientes de que las velocidades, incentivos, lenguajes y objetivos, a veces, pueden no converger y, tras la puesta en marcha de Evaluación 4.0, este ha sido un hecho que hemos experimentado. De todos modos, el conjunto de aliados en el consorcio considera la experiencia obtenida con Evaluación 4.0 de vital importancia para alimentar positivamente posteriores pilotos de acercamiento entre la industria 4.0 y la comunidad educativa, aprender de esta y ayudar a perfeccionar y optimizar los procesos de planificación y diseño de futuros proyectos. De igual manera, entendemos que **para que el campo de la investigación y la innovación tecnológica pueda realmente impactar positivamente con sus avances en la tarea de la comunidad docente, se hacen necesarios una coordinación y un diálogo constantes entre los tomadores/as de decisiones del mundo educativo, profesores/as que ejerzan el papel de líderes y referentes con sus pares.** Por supuesto, **también se necesita a las familias para conocer en profundidad sus inquietudes, incertidumbres y, por encima de todo, sus ilusiones y motivaciones** de querer formar parte de procesos de innovación tecnológica en los que la IA (o cualquier otra tecnología) pueda mejorar potencialmente los procesos educativos de sus hijos/as.

Algo que también anticipamos en la redacción inicial del proyecto, que pudimos constatar al ejecutarlo y que fue llamativo para algunos de los miembros del consorcio (sobre todo para aquellos menos cercanos al contexto educativo en su día a día), fue **la gran carga de trabajo que recae sobre el profesorado** (en este caso, de educación secundaria). Entre otras quejas, esgrime:

- **Falta de tiempo para poder participar en proyectos experimentales, avanzados y de innovación educativa.**
- **Poco reconocimiento a la labor docente y su implicación en la mejora de los procesos de aprendizaje** del alumnado.
- **Distanciamiento pronunciado entre la investigación científica y las oportunidades educativas** y sus contextos.
- **Ausencia de un discurso común y una visión compartida de los retos a superar** por parte de investigadores/as y realidades docentes debido a sus distintas inquietudes.
- **Falta de incentivos vinculados a la participación e implicación en procesos de innovación e investigación educativas.**

A tenor de lo que hemos mencionado anteriormente acerca de la gran carga de trabajo que recae sobre el profesorado, a lo largo del proyecto se ha constatado un factor que ha sido reiteradamente destacado en nuestras reuniones con los diferentes aliados del consorcio. Inicialmente, la búsqueda de centros participantes, que creemos que fue un gran reto que se consiguió superar, provocó un retraso inicial de los plazos de presentación de evaluaciones por parte de un porcentaje relativamente alto del profesorado participante en el experimento. En cualquier caso, esta experiencia

afianza aún más la esperanza en soluciones como PAAS para aligerar esta gran carga sobre el colectivo docente.

En definitiva, consideramos que este proyecto ha sido muy valioso, tanto en cuanto nos ha ayudado:

ÁMBITO EDUCATIVO

- **A ser pioneros y abrir camino a nuevos retos y aprendizajes**, que serán compartidos en abierto en beneficio de toda la comunidad educativa.
- **A poner en el foco** de la investigación y del piloto experimental la educación y **la equidad, anteponiéndolas a los intereses del avance tecnológico.**
- **A crear redes, comunidad y sinergias entre actores provenientes de diferentes mundos**, realidades e industrias (educativa, tecnológica y científica).
- A testar y pilotar un **desarrollo tecnológico vanguardista y puntero en contextos educativos reales**, donde habitualmente no suele llegar.

ÁMBITO CIENTÍFICO Y TECNOLÓGICO

- **A abrir camino en procesos de innovación tecnológica vinculada a la IA y la personalización del aprendizaje.**
- **A mejorar un algoritmo que pueda ayudar y dar apoyo a docentes** en tareas operativas.
- A estudiar la posibilidad de seguir investigando algoritmos y **soluciones que apunten a otras situaciones, tales como el abandono estudiantil prematuro.**

Recomendaciones en clave de política pública para el uso de la IA en educación

La UNESCO publicó en el 2019 el [Consenso de Beijing sobre la inteligencia artificial y la educación](#), que considera la IA como una tecnología capaz de mejorar las capacidades humanas, por lo que se hizo necesario llegar a acuerdos y establecer un marco seguro para su utilización y desarrollo con miras a “una colaboración eficaz entre el ser humano y la máquina en la vida, el aprendizaje y el trabajo, y para el desarrollo sostenible”.

A partir de los acuerdos del Consenso de Beijing, se han establecido nuevos marcos regulatorios para el desarrollo ético y responsable de la tecnología basada en IA y su aplicación en el ámbito educativo. La educación es un entorno sensible que requiere una consideración cuidadosa para asegurar que los derechos de la infancia no sean violados y que se fomente una educación inclusiva, equitativa y centrada en el aprendizaje de por vida.

Con el fin de destacar la importancia de proteger los derechos de la infancia, UNICEF (2021) publicó una [guía con orientaciones políticas](#) para garantizar la seguridad de los niños/as en relación con el uso de la IA en la educación. La guía establece tres requisitos fundamentales:

- **Proteger los derechos y la privacidad** de los niños/as en todo momento.
- **Fomentar el acceso equitativo** y la inclusión de todos los niños/as, independientemente de su origen o condición.
- **Promover el aprendizaje de por vida** y la participación activa de los niños/as en su propio proceso educativo.

Estas recomendaciones son vitales para garantizar el uso seguro y responsable de la IA en la educación y para asegurar que los derechos de la infancia sean respetados en todo momento. En el 2021, el Consejo de Europa reconoció que el sector educativo y el cultural son sectores especialmente sensibles por su vinculación con los derechos y valores de las personas debido a sus campos de acción y su naturaleza. Por lo tanto, a través de una [resolución del Parlamento Europeo](#), se ofreció una serie de recomendaciones para esos sectores. En esta línea, recogemos aquí algunas de las recomendaciones que consideramos más relevantes [4]:

- La IA es una herramienta al servicio de las personas y del interés general que ofrece oportunidades para abordar los principales desafíos de la educación.
- El marco educativo debe aprovechar las oportunidades de la IA desde una perspectiva de implementación que garantice los derechos fundamentales.
- Los algoritmos y la IA deben integrar los principios éticos desde el proceso de diseño.
- La IA puede y debe promover un aprendizaje profundo de calidad, compatible, inclusivo y que respete y proteja la igualdad de género, el multilingüismo y el diálogo intercultural.
- Es necesario recopilar datos fiables, con seguridad, sistematización y transferibilidad, sin que la privacidad se vea afectada.
- La adquisición de competencias digitales por parte de toda la ciudadanía es una condición previa para la transformación digital en beneficio de todo el mundo.
- Los beneficios de la implementación de la IA no dependen exclusivamente de la tecnología, sino del uso pedagógico que los y las docentes sean capaces de implementar.
- No se debe emplear la IA en detrimento de la educación presencial, a fin de potenciar el desarrollo social del alumnado y las capacidades cooperativas.
- Los estados miembros deben llevar a cabo una inversión en equipamiento digital en las escuelas que garantice la igualdad de oportunidades en la implementación de la IA en educación.

- Las medidas adoptadas para la reducción de sesgos y para evitar la discriminación de colectivos en el desarrollo y cuidado de la implementación de la IA van encaminadas a favorecer el progreso tecnológico en este sector.

Dicho todo lo anterior, desde el grupo impulsor se nos ocurren una serie de acciones que ya podrían desarrollarse y que irían en la línea de los marcos regulatorios mencionados anteriormente y buscarían apuntalar algunos de los retos que este informe ha mencionado y que se han experimentado en el proyecto Evaluación 4.0:

- La creación de un consejo asesor en España, compuesto por perfiles diversos (educadores/as, ingenieros/as, investigadores/as, líderes de comunidades de familias y líderes de movimientos estudiantiles), centrado en el uso correcto de la IA en educación.
- La vinculación y el trabajo en red de los distintos observatorios de ética e IA actuales junto con la Administración pública de nuestro territorio.
- El apoyo y la orientación al cuerpo docente con guías y recomendaciones claras sobre el uso de las herramientas tecnológicas que se ofrecen desde el sector privado y criterios de elección de unas u otras, según sus necesidades y proyectos educativos (atendiendo a criterios de equidad, justicia social e inclusión).
- La valoración del posible desarrollo de soluciones tecnológicas públicas que implementen la IA, en el marco de una mejoría de los retos que se nos plantean en el mundo educativo actual.



Conclusiones y retos pendientes

A raíz de las dificultades que afrontan los modelos neuronales para crear narrativas lógicas de las cadenas de decisión que los llevan a una predicción final, se plantea una controversia sobre sus aplicaciones en temas sensibles (como la evaluación automática). O, dicho de otro modo, **no consta suficiente evidencia sobre los criterios elegidos en la toma de decisiones de la IA aplicada en procesos de evaluación** automática que nos permita tener una confianza plena en sus decisiones. En respuesta a esas dificultades, una nueva ola de académicos/as de diferentes áreas está volcándose en el desarrollo de herramientas de inteligencia artificial que permitan apoyar y enriquecer la tarea evaluativa del profesorado, a la vez que ofrecen resultados trazables y no sesgados por bases de datos de entrenamiento. Ejemplo de ello es el modelo que aquí hemos utilizado, PAAS, que emplea conceptos de teoría multiagente (concretamente, la noción de confianza y su dispersión en una red social) para calcular la nota que un o una docente pondría a un examen que no ha podido revisar.

Para ello, PAAS calcula la importancia que el profesor/a otorgaría a la opinión de otros coevaluadores/as (alumnos/as). En estudios anteriores, habíamos observado que PAAS era capaz de hacer predicciones con un margen de error del 10 % con solamente un 30 % de los ejercicios corregidos por el tutor/a. En este caso, **no hemos podido apreciar esta tendencia de forma robusta**, en parte porque **el tamaño de las muestras por clase ha sido entre 10 y 30 veces menor que en dichos experimentos**, con lo que el algoritmo no ha sido capaz de aprender con precisión el perfil evaluador del o la docente ni de los alumnos/as. En las clases en las que el número de evaluaciones por alumno ha sido de en torno a 2,5 o más, y en las que los y las docentes han verificado todos los ejercicios a su cargo, **sí se ha apreciado un descenso monótono del error de predicción a medida que se introducían nuevas evaluaciones del tutor/a**. Es decir, podemos intuir (que no afirmar) que con un mayor volumen de evaluaciones, prácticas y experimentos podría mejorarse la cercanía del algoritmo al perfil evaluador del o la docente. Por supuesto, el motivo por el que las evaluaciones del profesorado han sido escasas en todas las clases responde a las limitaciones de tiempo, que no nos permitieron plantear la tarea como una labor individual. Al crearse equipos de trabajo, el número de ejercicios a revisar por docente se redujo al número de grupos a su cargo. **Aunque esta decisión facilitó el desarrollo del experimento, también supuso un hándicap para el algoritmo**. A todo esto, hay que añadir la pérdida de un 30 % de las evaluaciones de docentes que se produjo durante el proceso de recogida de datos, desconocemos si por error humano o de la plataforma.

Los mejores resultados de PAAS se dieron en aquellas clases en las que los alumnos/as mostraron **diferencias** en sus evaluaciones con respecto a las del profesor/a **estable**, hecho que confirma que PAAS es de mayor utilidad en situaciones en las que se cumpla la hipótesis de confianza transitiva.

A fin de validar más robustamente nuestro algoritmo, es deseable que en iteraciones futuras del experimento **se cuente con un mayor número de evaluaciones por docente, así como con un mayor número de coevaluaciones por ejercicio.**

Nuevas líneas de trabajo y aspiraciones futuras

Creemos que uno de los retos del presente (que no del futuro) es **mantener la cercanía entre la innovación tecnológica y científica y el campo educativo**, que ofrece una esperanza de mejora de la sociedad, para intentar hacer un mundo mejor. Dicho esto, consideramos que proyectos como Evaluación 4.0 deberían ser un ejemplo de buena praxis para incentivar la aparición de más propuestas como esta. Proyectos donde sigamos experimentando, probando, cuestionando y, por supuesto, **mejorando y adaptando la tecnología a las necesidades reales de la educación**, para poder ofrecer mejores oportunidades educativas, más equitativas y justas, a nuestro alumnado.

Valoramos muy positivamente los lazos establecidos durante esta experiencia como un activo de gran valor para todo el consorcio, **y queremos hacer hincapié en la importancia de juntar bajo un mismo propósito y estrategia la investigación y la docencia.** Además, creemos que sería fructífero para todos los miembros del consorcio continuar fomentando la relación que hemos iniciado. **Esta relación mantiene un círculo de líderes educativos —y a sus docentes— muy cercanos a la innovación tecnológica**, que poco a poco **pueden generar el discurso y la confianza que ambos mundos necesitan.** De este modo, suponemos que introducirán de manera más simple, rápida y óptima estas mejoras en sus aulas, que podrán formar parte de otros pilotos o innovaciones, como algunas que ya se están estudiando tras este primer piloto.

En este sentido, una posible continuación del proyecto en esta dirección permitiría no solo probar las predicciones del modelo en clases de alumnos/as reales, sino también acceder a informantes clave antes del testeo (*i. e.*, durante la fase de desarrollo) que inspiren nuestros experimentos de simulación y perfilen las preguntas de investigación. Es decir, la herramienta debería poder ofrecer al profesorado recomendaciones referentes a la gestión de los grupos de estudiantes. **La idea es que, dada una clase de estudiantes que deseamos que trabajen de manera colaborativa, la herramienta sugiera qué perfiles se complementarían mejor para lograr dicho objetivo**, qué número y composición tendrían los equipos óptimos o qué tipo de rúbrica debería utilizar cada equipo, entre otros aspectos relevantes de la organización del aula.

En definitiva, se plantearía una herramienta para la optimización de la formación de grupos, buscando que estos aprendan el máximo posible. Para ello, se generaría un modelo virtual de la clase concreta sobre la que se fuera a actuar, teniendo en cuenta aspectos individuales de los y las estudiantes, tales como sus competencias en cooperación, su nivel de competitividad o individualismo, su capacidad de influir en los demás miembros del grupo o su nivel de destreza a la hora de resolver ciertas tareas específicas. Una vez generado ese *clon* virtual de la clase, se optimizarían los equipos, que el o la docente podría organizar teniendo en cuenta las predicciones del modelo de los procesos grupales que acaecerían

con mayor probabilidad, y se buscaría el sistema de evaluación que, según el algoritmo, garantizase una mayor motivación y cohesión de grupo y mejores resultados en cada equipo.

Agradecimientos

El proyecto Evaluación 4.0 nació con **la voluntad de mejorar la educación de hoy en día, intentando que la tecnología fuera una herramienta al servicio de dicho propósito.** Desde los inicios del proyecto, la mirada que se estableció por parte de todo el equipo impulsor quiso mantener un **equilibrio, a veces difícil, entre, por un lado, la curiosidad y la esperanza en un futuro mejor, usando las potencialidades que la IA puede ofrecer,** y, por otro lado, **la incertidumbre y la precaución ante los posibles riesgos** que esta herramienta pueda tener asociados.

En el camino, se unieron muchas personas, entidades y aliados a los que, en nombre de todo el equipo impulsor, queremos agradecer su colaboración, participación y ofrecimiento a formar parte de este proceso de innovación, en consonancia con nuestra mirada inicial. **Nos llena de orgullo pensar que dicho camino, hecho en comunidad, pueda sembrar una semilla en las bases de futuros pilotos e iniciativas que promuevan la unión de la ciencia, la innovación y la tecnología al servicio de los objetivos educativos,** con perspectivas justas y equitativas, a fin de ofrecer mejores oportunidades educativas.

Dicho esto, no quisiéramos finalizar este informe sin mencionar **nuestro agradecimiento a todas las familias de los y las estudiantes que han formado parte del piloto** y, por consiguiente, a los alumnos/as que han participado activamente en el mismo. Asimismo, en nombre de todos los socios del grupo impulsor (Fundació Bofill, IIIA-CSIC, GAIA, Kid's Cluster y RO-BOTICA), queremos concluir este informe **dando las gracias a todos y cada uno de los centros que nos abrieron sus puertas —así como al conjunto de docentes y gestores/as** que lo hicieron posible—, y que citamos a continuación:

- Colegio Academia Jardín (Valencia)
- Colegio Martí Sorolla (Valencia)
- Institut Sant Pol de Mar (Sant Pol de Mar)
- Institut Salvador Espriu (Salt)
- Institut Salvador Sunyer i Aimeric (Salt)
- Institut Vallvera (Salt)
- Col·legi Amor de Dios (Barcelona)
- Institut Escola Daniel Mangrané i Escardó (Tortosa)
- Institut la Mar de la Frau (Cambrils)

Anexos

EXPERIENCIA EN LOS CENTROS EDUCATIVOS

A continuación se comparten las observaciones aportadas por el equipo de profesionales que llevó a cabo los talleres de programación en las diferentes aulas de los centros que participaron en el piloto, con el fin de contextualizar la experiencia, la muestra y los datos recogidos para alimentar el algoritmo.

Grupo: 01 Sesión: 1/1	Nombre del centro: La Mar de la Frau	Nivel: 3.º A (ESO)
N.º de alumnos/as: 20 N.º de participantes: 18	Provincia: Tarragona	Población: Cambrils
Fecha: 01/12/2022	Facilitadora: Rocío Lara	

OBSERVACIONES

En general, el nivel del alumnado del centro es medio-alto, en referencia a su conocimiento de la temática, por lo que está habituado al uso de la tecnología. Durante la presentación del experimento, los y las estudiantes se mostraron receptivos y participativos, y aportaron ejemplos de situaciones en las que está presente la IA.

También habían realizado satisfactoriamente el test previo para la conformación de los grupos. Cuando se llevó a cabo “el sorteo” desde la herramienta, mostraron sorpresa por los emparejamientos, pero en seguida se reubicaron para iniciar la actividad.

En sus clases no habían trabajado con la placa micro:bit, pero sí que habían realizado programas con Scratch y manipulado placas Arduino.

Al inicio de la sesión, se trabajó con un solo grupo (3A), que realizó toda la experiencia con la misma facilitadora, y durante la segunda hora se unió el grupo 3C, que cambió de aula para llevar a cabo la evaluación. Ambos grupos trabajaron juntos, pero cada uno con sus equipos previamente definidos.



A pesar de indicarles que el proceso de coevaluación no debía tener en cuenta quiénes eran las personas evaluadas, y que para ello los grupos no incluían ningún nombre o referencia que pudiera identificar a los autores/as del producto a evaluar,

fue complicado mantener este anonimato porque algunos alumnos/as se preguntaban continuamente de quién era cada película. En otros equipos el proceso fue transparente y efectuaron sus valoraciones sin tener en cuenta quién las había realizado. Estos fueron la gran mayoría, sobre un 70% del total.

Los y las docentes responsables de los grupos no se involucraron directamente en la sesión, pero sí que estuvieron pendientes de que el seguimiento, y sobre todo la disciplina y la cooperación del alumnado, fueran los adecuados.

Grupo: 02 Sesión: 1/1	Nombre del centro: La Mar de la Frau	Nivel: 3.º B (ESO)
N.º de alumnos/as participantes: 23	Provincia: Tarragona	Población: Cambrils
Fecha: 01/12/2022	Facilitador: Lluís Molas	

OBSERVACIONES

Al principio de la sesión explicamos qué era la IA, así como sus beneficios. Se mostraron muy receptivos y participativos. La sesión se desarrolló con una pequeña incidencia por la que se crearon dos grupos y los usuarios entraron a realizar la actividad y evaluarla en uno de esos dos grupos creados. Los alumnos/as pudieron realizar la creación y enlazarla en la plataforma, y también pudieron evaluar gran parte de las creaciones. Ya habían hecho actividades de *stop motion*, lo cual les fue útil para entender mejor la actividad. Pudieron descargar las creaciones en las placas micro:bit antes de finalizar la sesión.

Grupo: 03 Sesión 1/2	Nombre del centro: La Mar de la Frau	Nivel: 3.º C (ESO)
N.º de alumnos: 20 N.º de participantes: 18	Provincia: Tarragona	Población: Cambrils
Fecha: 01/12/2022	Facilitador: Lluís Molas	

OBSERVACIONES

La actividad se desarrolló fluidamente y con normalidad. Al inicio se explicó qué es la IA y se hizo hincapié en la seguridad de los datos almacenados y la precaución que se debe tomar en los entornos virtuales. Cabe destacar la originalidad en las diferentes creaciones que realizaron en la actividad. Este grupo ya había hecho alguna práctica de *stop motion* en cursos anteriores. Pudieron subir la actividad a la plataforma y realizar la evaluación de gran parte de las prácticas de sus compañeros/as.

Grupo: 03 Sesión 2/2	Nombre del centro: La Mar de la Frau	Nivel: 3.º C (ESO)
N.º de alumnos/as: 20 N.º de participantes: 18	Provincia: Tarragona	Población: Cambrils
Fecha: 01/12/2022	Facilitadora: Rocío Lara	

OBSERVACIONES

Con este grupo solo se llevó a cabo la experiencia de evaluación sobre las producciones que habían realizado la hora anterior con el grupo 3B y el facilitador.

En este segundo grupo se observó un poco más de dispersión en el momento de llevar a cabo las evaluaciones, pero en principio consiguieron terminar todos los registros.



Grupo: 04 Sesión: 1/1	Nombre del centro: Institut Escola Daniel Mangrané	Nivel: 3.º A (ESO)
N.º de alumnos/as: 21 N.º de participantes: 17	Provincia: Tarragona	Población: Jesús
Fecha: 18/11/2022	Facilitadora: Rocío Lara	

OBSERVACIONES

Prácticamente todo el alumnado de este centro lleva trabajando junto desde los 6 años, ya que se trata de un instituto escuela, lo que quedó reflejado en el ambiente y en la dinámica organizativa a la hora de realizar la actividad.

También la implicación del equipo docente fue determinante para llegar con los cuestionarios para la generación de los grupos cumplimentados de forma satisfactoria y trabajar ágilmente la conformación de los grupos. A pesar de este profundo conocimiento del alumnado, uno de los docentes se mostró sorprendido con las agrupaciones resultantes. Al final de la experiencia, manifestó su interés por Eduteams, ya que pudo comprobar la buena sintonía que se generó en aquellos grupos que, como expresó en sus propias palabras, "nunca se me hubiera ocurrido ponerlos a trabajar juntos".

Todos los equipos manifestaron haber completado las evaluaciones e incluso se pudieron probar los proyectos sobre las placas físicas, lo que añadió un punto muy positivo al momento de valorar los proyectos de sus compañeros/as.



Grupo: 05 Sesión: 1/1	Nombre del centro: Institut Escola Daniel Mangrané	Nivel: 3.º B (ESO)
N.º de alumnos/as participantes: 20	Provincia: Tarragona	Población: Jesús
Fecha: 18/11/2022	Facilitador: Lluís Molas	
OBSERVACIONES		
<p>Previamente a la actividad, el grupo ya estaba bien preparado para poder llevarla a cabo con agilidad y garantías. Al realizar los grupos, se sorprendieron gratamente de que la herramienta hubiera puesto en el mismo grupo de trabajo a una pareja que son muy amigos desde hace muchos años (es un centro que mantiene los mismos grupos clase desde los 6 años).</p> <p>La actividad se desarrolló con normalidad y el apoyo del profesorado fue muy bueno. Ya habían trabajado la programación anteriormente, aunque de forma básica.</p>		

Grupo: 06 Sesión: 1/2	Nombre del centro: Institut Amor de Dios	Nivel: 4.º (ESO)
N.º de alumnos/as: 20 N.º de participantes: 16	Provincia: Barcelona	Población: Barcelona
Fecha: 16/12/2022	Facilitadora: Rocío Lara	
OBSERVACIONES		
<p>De igual forma que en otros centros, Eduteams despertó curiosidad en el docente por las agrupaciones generadas, y al final de la experiencia se mostró satisfecho con el trabajo que pudo realizar su alumnado.</p> <p>En esta primera sesión, se completaron las pautas de programación del ejercicio y se llevó a cabo la realización de las animaciones. A solicitud del docente, los alumnos/as destinaron 15 minutos de la siguiente sesión a acabar de perfeccionar</p>		

sus proyectos.



Grupo: 06 Sesión: 2/2	Nombre del centro: Institut Amor de Dïos	Nivel: 4.º (ESO)
N.º alumnos/as: 20 N.º de participantes: 16	Provincia: Barcelona	Población: Barcelona
Fecha: 20/12/2022	Facilitadora: Rocío Lara	

OBSERVACIONES

La segunda sesión se inició con la recuperación de los proyectos y la finalización de los mismos. Se descargaron en las tarjetas y se inició la ronda de coevaluación.

Los equipos intercambiaron las tarjetas programadas y acompañaron las evaluaciones revisando los enlaces del documento entregado en la plataforma.

Hubo un par de grupos que propusieron mejoras a las realizaciones evaluadas y complementaron la propuesta con ideas que obtuvieron de otros grupos. De todos los equipos con los que se llevó a cabo la experiencia, solo se observó esta práctica en estos dos.



Grupo: 07 Sesión: 1/1	Nombre del centro: Institut Vallvera	Nivel: 3.º (ESO)
N.º de alumnos/as: 20 N.º de participantes: 15	Provincia: Girona	Población: Salt
Fecha: 17/11/2022	Facilitadora: Rocío Lara	

OBSERVACIONES

Este centro está catalogado como de alta complejidad.

El equipo docente del centro estuvo muy implicado en la actividad. La profesora a cargo del grupo estuvo presente todo el tiempo y brindó su ayuda no solo para vigilar la disciplina del grupo, que fue excelente, sino también para dar soporte cuando surgieron problemas de configuración o descarga en la placa con algunos programas. Este fue el primer centro en el que se llevó a cabo la experiencia y, por lo tanto, fue muy bien poder contar con su ayuda.

El grado de motivación del alumnado fue determinante para el desarrollo de la actividad, y la configuración de los grupos fue recibida con cierta expectación. Hubo sorpresa por parte de la profesora, como ocurrió en otros centros, por agrupaciones que no hubiera tenido en cuenta en otras circunstancias.

Los grupos completaron sus proyectos y las evaluaciones en el tiempo estipulado, lo que permitió que pudieran descargar las creaciones en la placa. Visto el resultado, se tomó como premisa intentar realizar este ejercicio en todas las actividades siguientes, si el tiempo lo permitía, previamente a la evaluación, ya que se observó que el nivel de dedicación era mayor cuando veían el proyecto funcionando en el dispositivo físico.



Grupo: 08 Sesión: 1/1	Nombre del centro: Institut Salvador Sunyer	Nivel: 3.º (ESO)
N.º de alumnos/as participantes: 13	Provincia: Girona	Población: Salt
Fecha: 14/12/2022	Facilitador: Lluís Molas	

OBSERVACIONES

La actividad se desarrolló con normalidad. Cabe destacar, como incidencia no relevante, que no se visualizaron correctamente algunos botones al subir la actividad a la plataforma debido a los dispositivos Chromebook del alumnado. Lo solventamos utilizando la presentación que llevaba el facilitador y pudimos identificar cuáles eran cada uno de los botones. Las parejas que no participaron en la actividad subiendo la película a la plataforma pudieron evaluar el resto de los proyectos utilizando una rúbrica impresa en papel. Les gustó mucho poder hacerlo de esta forma, porque no tenían la autorización firmada por parte de sus tutores legales.

Los alumnos/as mostraron mucho interés en realizar la actividad y poder colaborar con el proyecto. Unos días antes de la actividad, vio la luz Chat GPT, y fue un motivo para explicarlo y poder argumentar con más solidez la importancia de su acción e implicación en la coevaluación.

Grupo: 09 Sesión: 1/1	Nombre del centro: Institut Salvador Sunyer	Nivel: 4.º (ESO)
N.º de alumnos/as participantes: 15	Provincia: Girona	Población: Salt
Fecha: 14/12/2022	Facilitador: Lluís Molas	

OBSERVACIONES

La actividad se desarrolló con normalidad. Al haberse efectuado el día anterior la actividad con los alumnos/as de 3.º de ESO, los de 4.º ya estaban advertidos y recibieron la actividad con mucho entusiasmo. Al igual que en las sesiones con el alumnado de 3.º del mismo centro, algunos botones de la plataforma no se visualizaban correctamente, pero pudimos

solventarlo.

De igual forma, los alumnos/as mostraron mucho interés en realizar la actividad y poder colaborar con el proyecto. Tanto es así que el jefe de estudios del centro se acercó al aula para comprobar gratamente el comportamiento de una alumna que estaba dando muchos problemas en el centro por su actitud. Las creaciones realizadas fueron muy originales y los alumnos/as pudieron subir las a la plataforma y evaluar gran parte de ellas. También las pudieron descargar y enseñar al resto de la clase.

Como unos días antes de la actividad había visto la luz Chat GPT, se explicó su utilidad y proyección, lo que ayudó a concluir unas sesiones formidables para el centro.

Grupo: 10 Sesión: 1/2	Nombre del centro: Institut Salvador Espriu	Nivel: 3.º A (ESO)
N.º de alumnos/as participantes: 2	Provincia: Girona	Población: Salt
Fecha: 12/12/2022	Facilitador: Bernat	
OBSERVACIONES		
<p>Viendo que solo dos alumnas habían traído la autorización y que Eduteams no permitía hacer solo un equipo de dos personas, se optó por no excluir a los estudiantes 39 y 40. De este modo, formando equipo con esta suerte de "estudiantes fantasma", se pudo crear dos equipos, cada uno de ellos integrado por una estudiante real y un estudiante fantasma.</p> <p>A partir de ahí, la sesión se desarrolló con normalidad.</p>		

Grupo: 10 Sesión: 2/2	Nombre del centro: Institut Salvador Espriu	Nivel: 3.º A (ESO)
N.º de alumnos/as participantes: 2	Provincia: Girona	Población: Salt
Fecha: 14/12/2022	Facilitador: Bernat	
OBSERVACIONES		
<p>Como solo dos estudiantes habían traído la autorización, hubo tiempo más que suficiente para que se evaluaran entre ellas. Ambas estudiantes realizaron películas bien desarrolladas. Se reprodujeron las películas en micro:bit y, viendo que había tiempo, ambos grupos salieron a presentar su película ante sus compañeros/as. Al resto de alumnos/as se les dio la rúbrica impresa para que evaluaran, en el papel y sin que contase para la actividad, las exposiciones de sus compañeras.</p>		

Grupo: 11 Sesión: 1/2	Nombre del centro: Institut Salvador Espriu	Nivel: 3.º B (ESO)
N.º de alumnos/as participantes: 10	Provincia: Girona	Población: Salt
Fecha: 13/12/2022	Facilitador: Bernat	
OBSERVACIONES		
<p>Se trata de uno de los grupos del centro con mayor número de participantes. La sesión fue bien, e incluso algunos grupos llegaron a poner música a sus películas. Sin embargo, cuando tocó guardar la película en Eduteams para poder seguir trabajando el próximo día, hubo estudiantes que no pudieron acceder a la plataforma con su usuario y contraseña. En estos casos, guardaron el enlace al proyecto en su correo electrónico para poder seguir con el proyecto en la siguiente sesión.</p>		

Grupo: 11 Sesión: 2/2	Nombre del centro: Institut Salvador Espriu	Nivel: 3.º B (ESO)
N.º de alumnos/as participantes: 10	Provincia: Girona	Población: Salt
Fecha: 14/12/2022	Facilitador: Bernat	
<p>OBSERVACIONES</p> <p>Me di cuenta de que el motivo por el que en la sesión anterior algunos participantes no habían podido acceder a la plataforma se debía a que la docente había cometido un error en la formación de grupos y en la exclusión de participantes. Por este motivo, en algunos casos el usuario no coincidía con la identidad del o la estudiante, ya que para poder ingresar en la plataforma y evaluar a sus compañeros/as, tuvieron que usar otro. Asimismo, la profesora abandonó el aula prácticamente toda la hora, quedándome yo solo con todo el grupo. Por todo ello, fue una sesión algo caótica. No obstante, todos los grupos colgaron su película y evaluaron a los demás grupos.</p>		

Grupo: 12 Sesión: 1/2	Nombre del centro: Institut Salvador Espriu	Nivel: 3.º C (ESO)
N.º de alumnos/as participantes: 5	Provincia: Girona	Población: Salt
Fecha: 12/12/2022	Facilitador: Bernat	
<p>OBSERVACIONES</p> <p>La actividad se desarrolló con normalidad. Todo fue bien. Algunos grupos pusieron música a sus proyectos y funciones en los botones A y B.</p>		

Grupo: 12 Sesión: 2/2	Nombre del centro: Institut Salvador Espriu	Nivel: 3.º C (ESO)
N.º de alumnos/as participantes: 5	Provincia: Girona	Población: Salt
Fecha: 13/12/2022	Facilitador: Bernat	
<p>OBSERVACIONES</p> <p>Ambos grupos entregaron sus proyectos y evaluaron. Se reprodujeron las películas en micro:bit y, viendo que había tiempo, ambos grupos incluso salieron a exponer sus películas ante sus compañeros/as, quienes contaban con la rúbrica impresa en papel para que, sin que contase para el experimento, los evaluaran.</p>		

Grupo: 13 Sesión: 1/2	Nombre del centro: Institut Salvador Espriu	Nivel: 3.º D (ESO)
N.º de alumnos/as participantes: 11	Provincia: Girona	Población: Salt
Fecha: 13/12/2022	Facilitador: Bernat	

OBSERVACIONES

Se trata del grupo con mayor número de participantes del centro. La sesión fue según lo previsto. Algunos grupos pusieron música a sus proyectos y funciones en los botones A y B.

Grupo: 13 Sesión: 2/2	Nombre del centro: Institut Salvador Espriu	Nivel: 3.º D (ESO)
N.º de alumnos/as participantes: 11	Provincia: Girona	Población: Salt
Fecha: 15/12/2022	Facilitador: Bernat	

OBSERVACIONES

Se cumplió con lo previsto: se reprodujeron los proyectos en micro:bit, los grupos colgaron sus proyectos y evaluaron. Al haber traído la autorización solo la mitad de la clase, hubo momentos en que difícilmente podía estar pendiente de todos los alumnos/as que me llamaban, sobre todo de aquellos/as que no habían traído la autorización y que, al no poder evaluar, no tenían qué hacer una vez decían haber terminado con su película. En otros casos, cuando menos estudiantes habían traído la autorización, se había optado por permitir que algunos grupos saliesen a exponer su proyecto ante la clase, pero en este, al ser la mitad, yo tenía que estar muy pendiente de los que sí la habían traído.

Grupo: 14 Sesión: 1/2	Nombre del centro: Institut Salvador Espriu	Nivel: 3.º E (ESO)
N.º de alumnos/as participantes: 7	Provincia: Girona	Población: Salt
Fecha: 12/12/2022	Facilitador: Bernat	

OBSERVACIONES

El grupo no tenía efectuado el test de personalidad. La actividad se desarrolló de acuerdo con lo previsto. Incluso con algunos y algunas estudiantes se empezó a poner música y funciones en los botones A y B.

Grupo: 14 Sesión: 2/2	Nombre del centro: Institut Salvador Espriu	Nivel: 3.º E (ESO)
N.º de alumnos/as participantes: 7	Provincia: Girona	Población: Salt
Fecha: 13/12/2022	Facilitador: Bernat	

OBSERVACIONES

Todo iba según lo previsto. No obstante, cuando aún quedaban unos 15 minutos para terminar la sesión, al tratarse de la última clase del día, muchos estudiantes empezaron a recoger ante la pasividad del docente. Les reiteré en varias ocasiones que era necesario terminar la evaluación antes de marcharse, pero fue en vano, ante excusas como "se me ha terminado la batería del ordenador". Se acordó con la Sra. Emma Fernández que llamaría a dichos alumnos/as en otro momento para que terminaran de evaluar. Por otro lado, prácticamente todas las películas se reprodujeron en micro:bit.

Grupo: 15 Sesión: 1/2	Nombre del centro: Sant Pol de Mar	Nivel: 4.º (ESO)
N.º de alumnos/as: 20 N.º participantes: 12	Provincia: Barcelona	Población: Sant Pol de Mar
Fecha: 28/11/2022	Facilitadora: Rocío Lara	

OBSERVACIONES

Instituto público localizado en el Maresme. La población estudiantil, según comentó la profesora, corresponde principalmente a un estatus medio-alto, por lo que tienen acceso a la tecnología y algunas veces es complicado que se sorprendan con alguna actividad.

De todos los grupos en los que he intervenido como facilitadora, ha sido el único en el que no se contó con ninguna alumna.

La actividad se realizó a primera hora de un lunes, con lo que costó un poco engancharlos en la dinámica. Una vez realizada la presentación de la actividad y los objetivos del proyecto, se dio paso a la realización de las animaciones. En esta parte, ya se encontraban más activos, el nivel de cooperación fue muy bueno y todos probaron sus proyectos en las placas.

La sesión finalizó con la entrega de los proyectos a la aplicación. En la segunda sesión se llevarían a cabo las evaluaciones.





Grupo: 15 Sesión: 2/2	Nombre del centro: Sant Pol de Mar	Nivel: 4.º (ESO)
N.º alumnos/as: 20 N.º participantes: 12	Provincia: Barcelona	Población: Sant Pol de Mar
Fecha: 30/11/2022	Facilitador: Lluís Molas	
OBSERVACIONES		

Grupo: 16 Sesión: 1/2	Nombre del centro: Colegio Martí Sorolla	Nivel: 3.º (ESO)
N.º de alumnos/as participantes: 40	Provincia: Valencia	Población: Valencia
Fecha: 29/11/2022	Facilitador: Jorge Pérez	
OBSERVACIONES		
<p>El alumnado completó correctamente el proyecto, a pesar de que en ocasiones, y debido al comportamiento de parte del alumnado, surgieron dificultades y se ralentizó el proceso.</p> <p>Trabajaron por parejas y evaluaron al menos dos proyectos de la clase.</p> <p>Respecto a la organización de la actividad, el profesorado de este centro sí que estableció inicialmente los grupos y asignó la actividad, lo que facilitó el comienzo de esta.</p>		

Grupo: 17 Sesión: 1/2	Nombre del centro: Colegio Martí Sorolla	Nivel: 3.º (ESO)
N.º de alumnos/as participantes: 21	Provincia: Valencia	Población: Valencia
Fecha: 29/11/2022	Facilitador: Jorge Pérez	
<p>OBSERVACIONES</p> <p>El alumnado completó correctamente el proyecto, sin dificultad y con buen comportamiento.</p> <p>Trabajaron por parejas y evaluaron al menos dos proyectos de la clase, y en algunos casos incluso más.</p> <p>Respecto a la organización de la actividad, el profesorado de este centro sí que estableció inicialmente los grupos y asignó la actividad, lo que facilitó el comienzo de esta.</p>		

Grupo: 18 Sesión: 1/2	Nombre del centro: Colegio Martí Sorolla	Nivel: 4.º (ESO) B
N.º de alumnos/as participantes: *	Provincia: Valencia	Población: Valencia
Fecha: 29/11/2022	Facilitador: Jorge Pérez	
<p>OBSERVACIONES</p> <p>El alumnado completó correctamente el proyecto, sin dificultad y con buen comportamiento, por lo que pudimos ampliar la clase con algo más de contenido.</p> <p>Trabajaron por parejas y evaluaron al menos dos proyectos de la clase, y en algunos casos todos.</p> <p>Respecto a la organización de la actividad, el profesorado de este centro sí que estableció inicialmente los grupos y asignó la actividad, lo que facilitó el comienzo de esta.</p>		

Grupo: 19 Sesión: 1/2	Nombre del centro: Colegio Martí Sorolla	Nivel: 4.º (ESO)
N.º de alumnos/as participantes: *	Provincia: Valencia	Población: Valencia
Fecha: 29/11/2022	Facilitador: Jorge Pérez	

OBSERVACIONES

El alumnado completó correctamente el proyecto, sin dificultad y con un comportamiento correcto.

Trabajaron por parejas y evaluaron al menos dos proyectos de la clase, y en algunos casos incluso más.

Respecto a la organización de la actividad, el profesorado de este centro sí que estableció inicialmente los grupos y asignó la actividad, lo que facilitó el comienzo de esta.

Grupo: 20 Sesión: 1/2	Nombre del centro: Colegio Academia Jardín	Nivel: 3.º (ESO)
N.º de alumnos/as participantes: 40	Provincia: Valencia	Población: Valencia
Fecha: 29/11/2022	Facilitador: Jorge Pérez	

OBSERVACIONES

El alumnado desarrolló correctamente el proyecto, completando sin dificultad los diferentes pasos establecidos en la rúbrica de la actividad. Su comportamiento fue correcto.

Trabajaron por parejas y evaluaron al menos dos proyectos de la clase.

Respecto a la organización de la actividad, el profesorado no estableció inicialmente los grupos ni asignó la actividad. Lo hicimos nosotros al ver que no se había completado este paso previo al inicio del proyecto.

Grupo: 21 Sesión: 1/2	Nombre del centro: Colegio Academia Jardín	Nivel: 4.º (ESO)
N.º de alumnos/as participantes: 40	Provincia: Valencia	Población: Valencia
Fecha: 29/11/2022	Facilitador: Jorge Pérez	

OBSERVACIONES

El alumnado desarrolló correctamente el proyecto, completando sin dificultad los diferentes pasos establecidos en la rúbrica de la actividad. Su comportamiento fue correcto.

Trabajaron por parejas y evaluaron al menos dos proyectos de la clase.

Respecto a la organización de la actividad, el profesorado no estableció inicialmente los grupos ni asignó la actividad. Lo hicimos nosotros al ver que no se había completado este paso previo al inicio del proyecto.

Referencias

[1] Gutierrez, P., Osman, N., Roig, C., Sierra, C. (2016, enero). "Personalised automated assessments". En *Proceedings of the International Joint Conference on Autonomous Agents and Multiagent Systems*, AAMAS.

[2] Lissak, G. (2018). "Adverse physiological and psychological effects of screen time on children and adolescents: Literature review and case study". *Environmental Research*, 164, 149-157.

[3] Hernando Calvo, A., Muncio Zúñiga, A., Vázquez Gutiérrez, A., Gardó Huerta, H., Martínez Romero, H. (2022). Fundació Bofill. *Los algoritmos a examen: ¿Por qué IA en educación?*

https://fundaciobofill.cat/uploads/docs/q/z/z/dzp-e40_guia_digital_040123_algoritmos.pdf

[4] Muncio, A., Vázquez, A. (2023). *Política pública i intel·ligència artificial, els deu principis bàsics pels quals cal promoure'n una regulació de l'ús en educació*. Equitat Digital.

<https://equitatdigital.cat/intel·ligencia-artificial-i-politica-publica/>

EVALUACIÓN 4.0

