

Diseño y validación de una herramienta para el análisis y predicción de la innovación metodológica en centros de educación secundaria a través del aprendizaje automático

Fernando Cabeza¹ <https://orcid.org/0009-0008-3122-5460>
José L Díaz² <https://orcid.org/0000-0002-4677-0970>
Almudena Sánchez³ <https://orcid.org/0000-0002-4246-4132>
Julián Roa^{4*} <https://orcid.org/0000-0002-4017-3067>

¹ Facultad de Educación y Ciencias de la Salud. Dpto. de Educación y Tecnología. UDIMA, Madrid, España. (correo-e: FernandoCabezaSanchez.pro@gmail.com)

² Facultad de Educación y Ciencias de la Salud. Dpto. de Matemáticas. UDIMA, Madrid, España. (correo-e: joseluis.diaz.p@udima.es)

³ Facultad de Educación y Ciencias de la Salud. Dpto. de Matemáticas. UDIMA, Madrid, España. (correo-e: almudena.sanchez.s@udima.es)

⁴ Facultad de Educación y Ciencias de la Salud. Dpto. de Educación y Tecnología. UDIMA, Madrid, España. (correo-e: julian.roa@udima.es)

* Autor a quien debe ser dirigida la correspondencia.

Recibido Feb. 29, 2024; Aceptado Abr. 10, 2024; Versión final May. 8, 2024, Publicado Ago. 2024

Resumen

Los principales objetivos de esta investigación son proporcionar una descripción detallada de la tecnología de aprendizaje automático (ML, por sus siglas en inglés) aplicada a la medición de la innovación y diseñar un modelo que permita predecir el grado de innovación en una determinada institución. La tecnología ML carece de suposiciones o preconcepciones y es capaz de procesar una gran cantidad de datos y variables. Tras procesar los datos, se construye el modelo ML empleando variables asociadas al contexto educativo, se realiza el entrenamiento y se construye una web para predecir el grado de innovación de una institución educativa. El modelo proporciona una precisión del 66% en la predicción de la innovación y permite discutir la influencia de las variables analizadas a la hora de predecir el uso de metodologías activas en una institución. En conclusión, este enfoque puede abrir nuevas técnicas de análisis de datos apoyadas en ML que complementen los enfoques tradicionales basados en la estadística.

Palabras clave: análisis de datos; aprendizaje automático; innovación educativa; innovación metodológica; modelos predictivos

Design and validation of a tool for the analysis and prediction of methodological innovation in secondary education institutions through machine learning

Abstract

The primary objectives of this research study are to provide a detailed description of machine learning (ML) technology when applied to assessing innovation and to design a model that allows predicting an institution's degree of innovation. Machine learning technology lacks assumptions or preconceptions and is capable of processing a large amount of data and variables. After data processing, the ML model is built using variables associated with educational context, training is performed, and a web is built to predict the degree of innovation of an educational institution. This model provides an innovation accuracy prediction of 66% and allows assessing the influence of the variables analyzed when predicting the use of active methodologies at a given institution. In conclusion, this approach can open new data analysis techniques supported by ML that complement traditional statistically based approaches.

Keywords: data analysis; machine learning; educational innovation; methodological innovation; predictive models

INTRODUCCIÓN

En primer lugar y antes de comenzar con los detalles del presente artículo, señalamos cómo debe interpretarse el término *innovación educativa* en el contexto del trabajo que se presenta. Por innovación educativa interpretamos un conjunto de actividades en el entorno educativo que fomentan la interacción entre estudiantes como forma activa de aprendizaje. De esta forma, entendemos que el punto de partida consiste en una enseñanza tradicional centrada en la figura del docente y de los contenidos a impartir, y es desde este punto de partida sobre el que evaluamos la capacidad de una institución educativa para llevar a cabo propuestas activas, en nuestro caso, consideradas como innovadoras. El modelo basado en Aprendizaje automático y, en particular, en Redes Neuronales que se plantea en este trabajo es capaz de establecer un criterio para saber si un centro tiene capacidad para poner en marcha metodologías activas de aprendizaje. Pongamos un ejemplo ilustrativo: Consideremos un centro educativo que sigue una metodología de aprendizaje tradicional. En base a la medición de una serie de variables contextuales y propias del centro (como se definirá más adelante en el trabajo), puede suceder que el modelo indique que el centro innova. En este caso, debemos interpretar que el centro tiene la suficiente capacidad en materia de recursos, formación del personal docente o incluso un entorno socioeconómico favorable (así como otras variables que se mencionarán) como para innovar y poner en marcha metodologías activas de aprendizaje. De esta forma, una puesta en práctica de las políticas correctas a nivel centro hacia propuestas pedagógicas activas, y alejadas del enfoque tradicional, debería culminar en una implantación exitosa de las mismas.

Antes de comenzar con la descripción de nuestro trabajo, nos gustaría comentar algunos aspectos educativos contextuales: El hecho de garantizar que todas las personas tengan las mismas oportunidades para completar sus estudios es una tarea difícil. Muchos sistemas educativos de distintos países tienen dificultades a la hora de atajar este problema estando, además, inmersos en un conjunto de leyes educativas que tratan de potenciar la capacidad del ecosistema educativo para atender a todas las necesidades individuales de cada tipo de alumnado. El hecho de educar, por y para todos, implica prestar atención a una realidad universalmente diversa, rica en culturas y pensamientos, que requiere del sistema educativo una prueba de su capacidad para proponer metodologías amplias, flexibles y capaces de innovar mientras se adaptan. La puesta en práctica de tales procesos no sólo depende de leyes educativas promovidas por las instituciones políticas, sino que existe consenso en considerar la relevancia de los centros educativos para adaptarse a los cambios e innovar en el transcurso dinámico del proceso educativo. Hacemos hincapié en el término dinámico, que hemos introducido apropiado para este fin, para señalar el carácter no estático de la innovación. Sin miedo a errar, podemos decir que una institución educativa innova para adaptarse a las nuevas demandas, a las nuevas realidades que se deben afrontar y para traer las nuevas tecnologías a las aulas. Admitiendo el carácter dinámico de esas demandas y realidades, con todas sus complejidades, encontramos que la innovación puede concebirse como el conjunto de respuestas que desde las instituciones educativas se proporcionan y cuya puesta en práctica actúa como fuente clave de conocimientos.

Desde la perspectiva mencionada, resulta pertinente cuestionar la capacidad de adaptación que posee una institución educativa específica ante el entorno complejo en el que se halla. Expresado de otra manera, se trata de reflexionar sobre la habilidad de la institución para innovar y sobre cómo determinamos un indicador actual de dicha innovación, que debería servir como un punto de partida y diagnóstico. No podemos afirmar con certeza que exista un consenso concreto sobre cómo medir la capacidad para innovar de una determinada institución. No obstante, según la OECD (2021), se puede prestar atención a cinco dimensiones básicas cuya correcta caracterización, y puesta en práctica concreta, permite establecer unas mínimas líneas hacia donde enfocar la medida de la innovación en una institución (y, a la vez, establecen zonas donde potenciar posibles cambios que impliquen innovaciones constantes): 1) Cultura favorable a la innovación; 2) Conocimientos, capacidades y motivación; 3) Gestión de la innovación; 4) Identificación de recursos e impulsores de la innovación; 5) Capital humano en la educación. Por otro lado, cabe destacar la importancia del desarrollo de una cultura favorable a la innovación en la institución escolar, la cual debe apostar por el fomento de un cierto nivel de libertad y autonomía en el cuerpo docente, aspectos que permitirán el desarrollo de entornos co-creativos (Shalley y Gilson, 2004).

A lo largo del transcurso escolar, se generan una gran cantidad de datos que pueden ser útiles para multitud de propósitos como bien muestra el Ministerio de Educación y Formación Profesional (MEFP, 2021). En particular, si se utilizan los datos adecuados se podrá analizar con más detalle la capacidad de una institución para innovar, así como establecer un marco de referencia sobre el grado de innovación en una institución. Parece conveniente preguntarse si es posible encontrar una manera sencilla de determinar qué condiciones son las idóneas para desarrollar propuestas innovadoras en un determinado centro educativo. Por otro lado, el conocimiento concreto del estado, y su capacidad, en una institución de cara a innovar es interesante para el capital humano de dicha institución, ya que, en base a ello, el profesorado podrá particularizar sus planes de formación para aprovechar al máximo las capacidades del centro medido.

En la actualidad existen diferentes herramientas empleadas para medir la capacidad de innovar que tiene una institución educativa. Algunas de las más famosas se encuentran recogidas en OECD (2021), tales como SELFIE, AGITIC, eLEMER. Todas ellas están basadas en cuestionarios validados por una amplia red de expertos y tratan principalmente de medir la capacidad de las instituciones para innovar con aprendizajes activos y el uso de la tecnología. El empleo efectivo de este tipo de herramientas se encuentra fuertemente influido por el uso de métricas e indicadores adecuados, los cuales pueden resultar algo subjetivos y su incorrecta interpretación puede poner en riesgo la conclusión sobre si una institución innova o no. Encontramos aquí un potencial aspecto de mejora en la forma de medir la innovación. Tratamos, por tanto, en este trabajo de proporcionar una forma diferente y relevante de medir la innovación en educación eliminando en la medida de lo posible los sesgos en la selección de métricas e indicadores. Para ello, se va a utilizar una de las herramientas tecnológicas más potentes de la actualidad, como es la inteligencia artificial, en particular la técnica *aprendizaje automático* (ML en lo sucesivo). Este enfoque ya ha sido explorado por otros autores como Mosquera et al. (2018) en el ámbito educativo y supone un enfoque alternativo a la técnica basada en cuestionarios.

De acuerdo con Sarker (2021), el *aprendizaje automático* se ha consolidado como una herramienta fundamental para abordar una amplia gama de problemas prácticos, y su versatilidad la convierte en una tecnología ideal para el manejo de datos, recopilación de información y extracción de ideas sobre la innovación potencial en un centro educativo. Las ventajas de aplicar la tecnología ML para la evaluación de innovaciones en instituciones educativas, en comparación con los enfoques convencionales de encuestas, son notables: 1) La tecnología ML carece de suposiciones o preconcepciones que podrían limitar el alcance de su análisis. Esto contrasta con los cuestionarios temáticos, donde las preguntas suelen estar influenciadas por las intenciones y expectativas de los investigadores; 2) El ML tiene la capacidad de procesar y analizar una gran cantidad de variables, incluso aquellas que a primera vista parecen no tener relación. Esto le permite descubrir conexiones y patrones significativos que pueden pasar desapercibidos en métodos de análisis más tradicionales. Desde las perspectivas mencionadas, nos proponemos los siguientes objetivos en el presente artículo: 1) Proporcionar una descripción detallada de la tecnología ML aplicada a la medición de la innovación; 2) Diseñar un modelo que nos permita predecir el grado de innovación en una determinada institución educativa en función a unas variables objeto a estudio.

MÉTODO

Esta sección se compone de varias subsecciones que tratan de ilustrar el protocolo metodológico seguido. Primero se introduce un contexto proporcionando al lector una breve explicación del conjunto de datos con el que se va a trabajar. En la siguiente parte se procede con la limpieza y acondicionamiento de los datos para su posterior uso. A continuación, se lleva a cabo una codificación de estos. Por último, se construye el modelo en base a principios de *aprendizaje automático* posteriormente se realiza el entrenamiento del modelo y se construye una web para que cualquier persona interesada pueda predecir el grado de innovación de una institución educativa. Adicionalmente, cabe destacar que el lenguaje de programación que se va a utilizar para implementar la metodología es Python al igual que en otros trabajos similares como el de Contreras *et al.* (2020). Según Raschka et al. (2020) actualmente es uno de los lenguajes de programación más potentes para la computación científica, la ciencia de datos y el aprendizaje automático, lo que permite aumentar el rendimiento y la productividad al hacer uso de librerías de bajo y alto nivel de calidad.

Explicación de los datos

El conjunto de datos, empleados para la investigación, ha sido construido por el grupo de investigación del Observatorio de Innovación Educativa (OIE) de la Universidad a Distancia de Madrid (UDIMA). El Observatorio se creó en el marco de los estudios de grado y máster que ofrece la UDIMA y con el objetivo de promover la actividad reflexiva de los estudiantes de último curso en la fase de prácticum. En particular, el Observatorio se ha centrado en aquellos estudiantes que han hecho su fase final de prácticas en un centro educativo, constituyéndose así en observadores imparciales de la realidad que viven en su etapa de prácticas. Ellos pueden analizar críticamente la realidad observada gracias a la formación recibida aún reciente. La implementación se mide mediante el uso de un cuestionario que recoge información sobre múltiples variables educativas observadas de forma imparcial durante los meses de estancia en centros educativos oficiales. Para asegurar la relevancia de las escalas en el cuestionario, se llevó a cabo un análisis basado en el coeficiente de competencia experta según Cabero y Barroso (2013), identificando cuatro áreas clave: Metodología, Organización, Recursos y Evaluación. Además, se llevó a cabo una revisión del instrumento por pares expertos en el ámbito educativo. El conjunto de expertos tiene una media de más de diez años en el sector educativo, además de ser reconocidos investigadores. Durante esta fase de validación, el grupo de expertos ha hecho propuestas de mejora, tanto en la formulación de las preguntas, como de las posibles respuestas en relación con los objetivos a medir. Las mejoras propuestas han sido implementadas. De esta forma, se han producido varios ciclos de mejora que han culminado en un cuestionario validado por expertos.

Adicionalmente se ha realizado un análisis de la fiabilidad y la validez mediante el estudio de la consistencia interna de los ítems y las componentes principales que ha permitido obtener garantías satisfactorias de validez y fiabilidad para conocer el estado de la innovación educativa en centros de enseñanza. En cuanto a las características psicométricas del cuestionario, se evaluaron la fiabilidad y la validez a través de la consistencia interna de los ítems y un análisis de componentes principales. Los resultados mostraron una buena fiabilidad ($\alpha = .75$) y una estructura factorial que explica la variación en las respuestas. Estos hallazgos psicométricos sugieren que el instrumento es confiable y válido para analizar la implementación de metodologías activas en el entorno educativo. Asimismo, cabe señalar que la obtención de datos a partir de cuestionarios en trabajos de ML en educación es avalada en otros trabajos como el de Sghir *et al.* (2023).

Variables explicativas/exógenas

El estudio se basa en una muestra recogida por estudiantes del Máster del Profesorado que cursaron la asignatura de Prácticum en sus respectivas especialidades. Esta muestra es relevante porque la asignatura de Prácticum incluye prácticas presenciales en centros educativos, lo que permite a los estudiantes ser observadores directos del uso de distintas metodologías innovadoras en el aula. Los registros recopilados se extienden desde el primer semestre del curso 2019/2020 hasta el segundo semestre del curso 2021/2022, con un total de 1930 entradas al formulario. Para esta investigación se van a tomar una serie de variables exógenas que se dividen en dos grupos como se indica a continuación y cuyas características vienen dadas de acuerdo a las siguientes descripciones: 1) Un grupo de variables que se pueden obtener a priori, es decir en base a los datos que se aportaron desde el Observatorio de la Innovación Educativa y que están relacionadas con medidas clave del propio sistema educativo. En concreto, se refieren a las siguientes variables: 1.1) Especialidad: área de conocimiento que imparte el tutor de prácticas en los centros educativos donde el estudiante ha realizado sus prácticas, las cuales son: Matemáticas, Educación física, Biología y geología, Lengua castellana y literatura, Lengua extranjera-Inglés, Geografía e historia, Formación profesional y Sin especialidad; 1.2) Edad: se compone por tres categorías y hace referencia a la edad del alumno en fase de prácticas de la UDIMA. Las categorías son las siguientes: menos de 25 años, entre 25 y 40 años (ambos inclusive) y más de 40 años; 1.3) Género: sexo del alumno en prácticas (masculino y femenino); 1.4) Comunidad autónoma: comunidad autónoma del territorio español (17 comunidades autónomas) más la opción de realizar las prácticas en un territorio no español (centros del Ministerio de educación y Formación Profesional - Gobierno de España); 1.5) Horario: horario del centro, se compone de seis categorías: Jornada de mañana o de tarde; Jornada continua; Jornada partida; Jornada de mañana y tarde; Jornada mixta; Jornada nocturna; 1.6) Tipo de centro (titularidad): El tipo de centro según su titularidad tiene tres opciones: público, privado, concertado; 1.7) Tipo de centro (según el alumnado): El tipo de centro, según el tipo de alumnado, puede ser: Centro de integración preferente auditiva; Centro de integración preferente motora; Centro de integración preferente para personas con Espectro del Trastorno Autista; Centro Ordinario; Centro Ordinario con inclusión/de integración/con apoyos; Otro; 1.8) Entorno demográfico: rural y urbano; 1.9) Entorno Demográfico: Se refiere al tamaño de la población donde se encuentra ubicado el centro escolar. Tiene cinco categorías: Menos de 5000 habitantes; Entre 5000 y 20000 habitantes; Entre 20000 y 50000 habitantes; Entre 50000 y 200000 habitantes; Entre 200000 y 1000000 habitantes; Más de 1000000 de habitantes; 1.10) Edad del tutor de prácticas: Las categorías son las siguientes: menos de 25 años, entre 25 y 40 años (ambos inclusive) y más de 40 años; 1.11) Años de experiencia del tutor: está formado por tres categorías: Entre 1 y 5 años (ambos inclusive); Entre 6 y 10 años (ambos inclusive); Más de 10 años; 1.12) Formación pedagógica del tutor: CAP; Máster de Formación del profesorado; Ninguna; Otra.; 1.13) Tipos de espacio: Lugar donde se desarrollan las clases. Las categorías son las siguientes: Aulario; Aula materia; Aula de estructura variable por agrupamiento; Aula abierta; Otra. 2) Otro grupo de variables que se utilizarán para construir la variable final a predecir y que tienen que ver con algunas de las metodologías de innovación más frecuentemente observadas de acuerdo a Díaz *et al.* (2023), son: Aprendizaje cooperativo; Gamificación; Flipped classroom; Aprendizaje basado en problemas; Aprendizaje basado en proyectos; Aprendizaje servicio.

Para todas y cada una de dichas metodologías, contábamos con cinco opciones de respuesta que son las siguientes: 1) No he visto que el trabajo en grupo forme parte de las metodologías empleadas; 2) He visto cómo se trabajaba en grupo pero sin cumplir alguna de las características de 'la metodología'; 3) He visto cómo se utiliza 'la metodología' de forma puntual (<25%); 4) He visto cómo se utiliza 'la metodología' en un número importante de sesiones (>25% y <75%); 5) He visto cómo se utiliza 'la metodología' en la mayoría de las sesiones (>75%). Donde 'la metodología' determina cualquiera de las metodologías sobre la cual se pregunta. Es relevante destacar que la formación recibida por parte de los estudiantes en el máster de formación del profesorado les capacita para la identificación de metodologías activas de acuerdo con Díaz *et al.* (2023). De esta manera, aseguramos que los observadores tienen la capacitación adecuada, así como la imparcialidad requerida al no formar parte de la institución educativa como profesorado permanente.

En el proceso de preparación de datos para el análisis, se lleva a cabo inicialmente una etapa de filtrado, reduciendo el conjunto de datos masivo a un subconjunto más manejable y relevante. Este paso es crucial para mejorar la eficiencia y la efectividad del análisis posterior. Una vez realizado el filtrado, se procede a examinar la naturaleza de las variables presentes en el conjunto de datos. Es esencial identificar el tipo de datos que cada variable representa, ya sean numéricos, categóricos o textuales, para poder aplicar las técnicas de análisis adecuadas.

Una vez que los estudiantes de prácticas han rellenado el cuestionario, comienza la etapa de preparación de los registros recogidos. Esta etapa es crucial debido a que la calidad y estructura de los datos influyen directamente en la capacidad de aprendizaje y precisión del modelo. Además, la codificación de datos en esta etapa es un paso fundamental que transforma los datos crudos en un formato estructurado y comprensible por los modelos computacionales (Hale, 2021). Desde este punto de vista y de forma concreta, se llevaron a cabo los siguientes pasos: 1) Verificación que no hay registros duplicados. En su caso, son eliminados; 2) Dado que un estudiante puede actuar como observador en aulas de diferentes niveles educativos y con el fin de realizar el estudio en cada uno de los contextos, se reestructura el conjunto de datos recogidos de modo que cada registro presenta las respuestas de un estudiante en un aula concreta. Así pues, el número de registros por aula aumentará en función del número de aulas que haya observado un estudiante pudiendo ser un número variable, es decir, un estudiante puede haber observado dos aulas (3º y 4º ESO) mientras que otro puede haber observado únicamente un aula (1º BACH); 3) Teniendo en cuenta la tipología de las variables, las cuales son en su mayoría categóricas, se lleva a cabo un proceso de codificación de las variables con el fin de facilitar el desarrollo de los procesos estadísticos requeridos con asignaciones numéricas a cada una de las opciones de respuesta.

Por ejemplo, en el caso de variables nominales, como puede ser el tipo de centro será codificado con el número 1 para centros privados, 2 para centros concertados y 3 para centros públicos; en el caso de variables de carácter ordinal, como puede ser las asociadas al uso de cada una de las metodologías objeto a estudio, la codificación será de un valor numérico de acuerdo al porcentaje de uso de dicha metodología, esto es, se asignará el valor 1 a la opción de respuesta “No he visto que el trabajo en grupo forme parte de las metodologías empleadas”; 2 a la respuesta “He visto cómo se trabajaba en grupo pero sin cumplir alguna de las características de ‘la metodología’”; 3 a “He visto cómo se utiliza ‘la metodología’ de forma puntual (<25%)”; 4 identificará “He visto cómo se utiliza ‘la metodología’ en un número importante de sesiones (>25% y <75%)”; 5 determinará “He visto cómo se utiliza ‘la metodología’ en la mayoría de las sesiones (>75%)”.

Una vez realizados estos pasos, el conjunto de datos se encuentra preparado para poder llevar a cabo los procesos estadísticos correspondientes. Este proceso de preparación garantiza la calidad y la fiabilidad del análisis subsiguiente, proporcionando una base sólida para la extracción de resultados significativos a partir de los datos (Witten et al., 2016).

Codificación de datos– Variable explicada/endógena

La codificación de las variables recogidas, principalmente, las asociadas al uso de cada una de las metodologías según el porcentaje de uso indicado en el apartado anterior con números naturales, servirá para la creación de una nueva variable que compute el grado de innovación que se ha observado en los centros, y que se denominará “*Multiclase*”. Dicha variable será el valor máximo registrado asociado al conjunto de cada una de las metodologías, esto es, si en un aula concreta (registro), un estudiante ha indicado que ha observado todas las metodologías, si bien, en una de ellas ha marcado el máximo porcentaje de uso, entonces la variable *Multiclase* tendrá ese valor máximo indicado. Para poder medir la presencia efectiva de innovación dentro del contexto educativo analizado, además, se introduce una nueva variable en el conjunto de datos, denominada “*Clase*”. Esta variable actúa como un indicador binario, asignando un valor de 0 para representar la ausencia de innovación y un valor de 1 para indicar su presencia. La asignación de estos valores se basa en un criterio definido por la columna “*Multiclase*”.

La columna “*Multiclase*” refleja la frecuencia de observación de sesiones con metodologías innovadoras. Se establece un umbral en el valor 4, el cual se ha determinado como el punto de corte crítico. Este umbral corresponde al percentil 25, lo que indica que en al menos el 25 % de las sesiones observadas se reportó la implementación de prácticas innovadoras. De esta manera, si el valor de la columna “*Multiclase*” es igual o superior a 4, la nueva columna “*Clase*” se etiqueta con un 1, indicando la presencia de innovación. Si el valor es inferior a 4, se asigna un 0, denotando la ausencia de innovación. Una vez que se ha construido la columna “*Clase*” y se han asignado las etiquetas correspondientes, se procede a eliminar las variables auxiliares que se utilizaron para este fin. Esto se hace para simplificar el conjunto de datos y centrar el análisis en la variable objetivo, “*Clase*”, eliminando así cualquier redundancia o información superflua que podría desviar la atención del objetivo principal del estudio. La técnica expuesta es similar a la de otros estudios en educación como los de Alyahyan et al. (2020).

Construcción y entrenamiento del modelo

Para abordar el desafío de medir y evaluar la innovación educativa dentro de una institución, es crucial identificar y definir una variable objetivo. En este estudio y como ya se ha mencionado, esta variable se representa en forma de datos binarios, tomando valores de 0 o 1, los cuales reflejan la ausencia o presencia de innovación, respectivamente. Con el fin de desarrollar un modelo predictivo, el conjunto de datos se segmenta en dos partes: un conjunto de entrenamiento y otro de prueba. El conjunto de entrenamiento comprende el 90% del total (1737 registros), mientras que el conjunto de prueba representa el 10% restante (193 registros). Esta partición se realiza de forma aleatoria, empleando un valor de semilla establecido en 7.

El análisis de datos, se lleva a cabo con el desarrollo de un modelo de aprendizaje automático (ML) basado en redes neuronales de tres capas, de las cuales, las dos primeras capas están compuestas por 50 neuronas cada una, y la tercera y última capa se constituye por un único perceptrón (neurona artificial) diseñado para la clasificación binaria. Esta aproximación ha sido analizada en Assi *et al.* (2019) y utilizada en trabajos similares como el de Contreras *et al.* (2020).

Para la optimización de los hiperparámetros (variables iniciales) del modelo, se emplea una técnica conocida como búsqueda en cuadrícula (grid search). Este método implica la evaluación exhaustiva de una serie de modelos candidatos, cada uno con diferentes combinaciones de hiperparámetros predefinidos. Este proceso permite seleccionar la configuración óptima basándose en el rendimiento del modelo durante la fase de entrenamiento. El modelo óptimo resultante, denominado "*bestModel*" en el código, demuestra una precisión media de 0.68 para el conjunto de datos de validación. Los hiperparámetros que conducen a este resultado son: un número de épocas igual a 36 (*bestEpochs*), un tamaño de lote de 100 (*bestBatchSize*) y un valor de división de validación de 0.1 (*bestValidationSplit*).

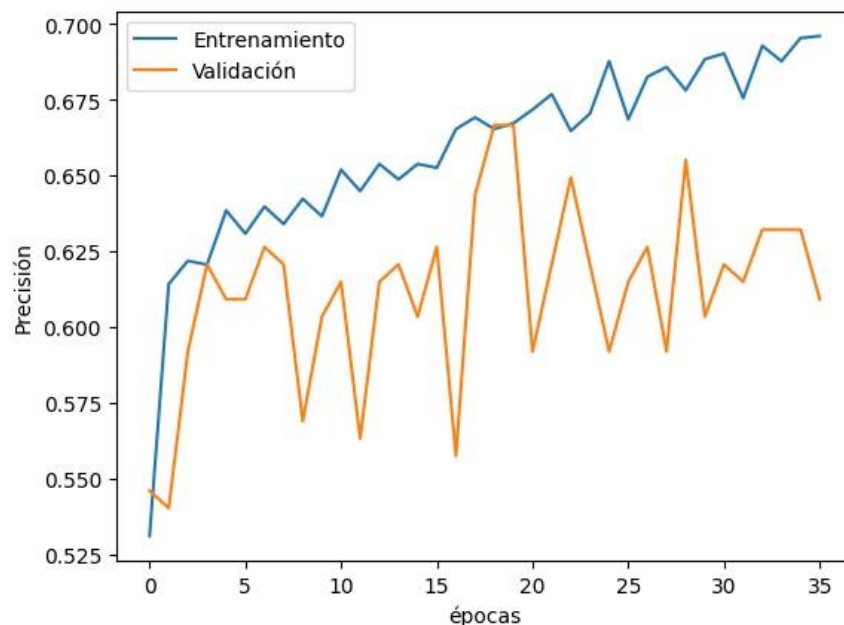


Fig 1: Evolución del entrenamiento del modelo

RESULTADOS

La evaluación de un modelo de clasificación requiere la construcción de una matriz de contingencia, la cual compara los valores reales con las predicciones realizadas por el modelo. Este paso es crucial para analizar las métricas relevantes que reflejan la eficacia del modelo. Es importante destacar que este análisis se lleva a cabo utilizando el conjunto de datos de prueba, que no han participado en la fase de entrenamiento del modelo. La creación de la matriz de contingencia es el preludeo para examinar detalladamente las métricas asociadas con la precisión del modelo (para más información, se recomienda recurrir a Tharwat, 2021). Para evaluar un modelo de clasificación lo que se debe hacer es construir la matriz de contingencia que relaciona la realidad con la predicción para posteriormente analizar sus métricas asociadas. Todo esto se realiza para los datos test, que no han intervenido en el proceso de entrenamiento. Se procede con la elaboración de las tablas de contingencia. Si se quiere indagar sobre las métricas que evalúan la precisión de un modelo se recomienda la lectura de Tharwat (2021).

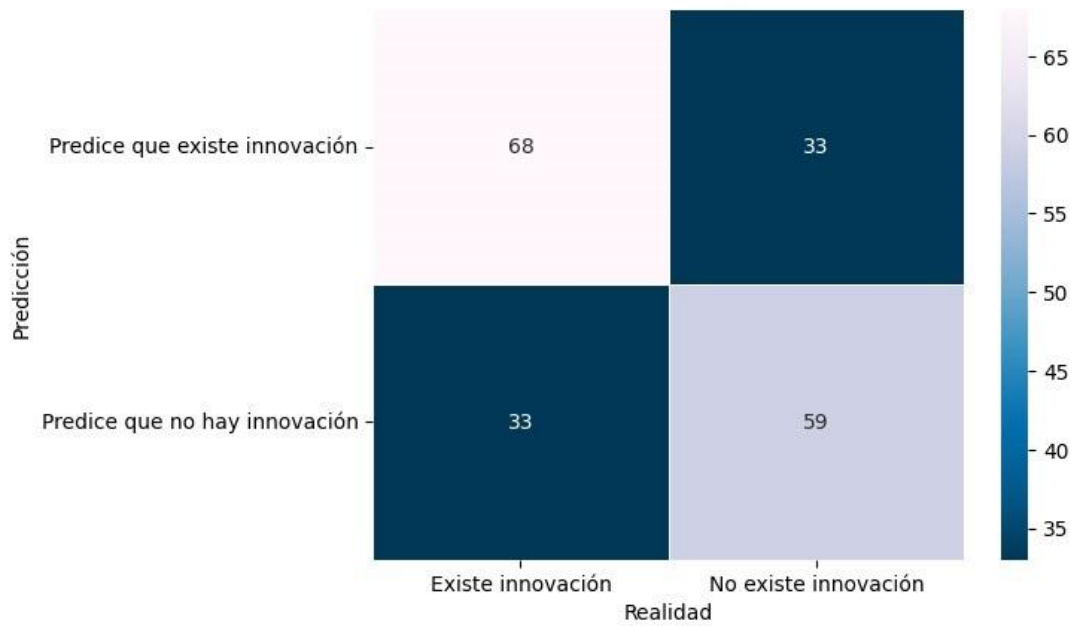


Fig. 2: Tabla de contingencia en bruto

La matriz de contingencia muestra que hay 68 casos en los que la innovación efectivamente ocurre y el modelo lo identifica de manera acertada. Sin embargo, existen 33 instancias en las que, a pesar de haber innovación, el modelo no logra reconocerla. Por otro lado, el modelo indica la presencia de innovación en 33 ocasiones donde, en realidad, no se da tal innovación. Finalmente, el modelo es capaz de identificar correctamente la ausencia de innovación en 59 casos.

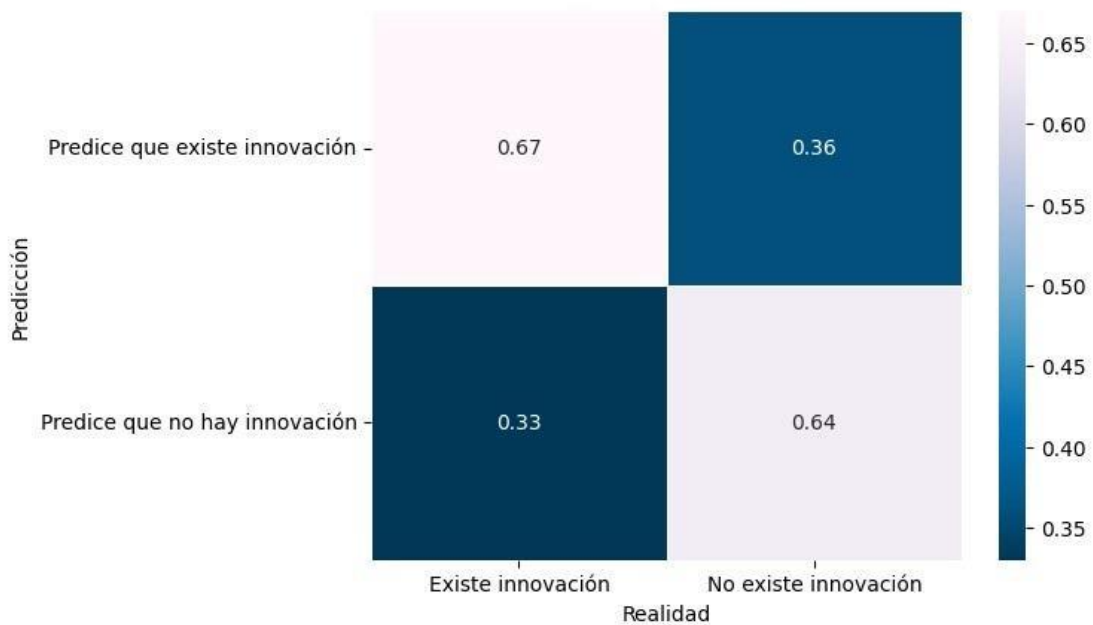


Fig. 3: Tabla de contingencia proporción por columnas

La sensibilidad del modelo, es decir, la proporción de medir correctamente la innovación, sabiendo que de verdad se innova es de 0'67. La especificidad, esto es, la proporción de medir que no hay innovación sabiendo que realmente no se innova es de 0'64. Claramente la curva ROC está por encima de la diagonal, lo que implica un grado de clasificación bastante mejor que la predicción aleatoria. Además, el área bajo la curva es de AUC = 0.69. Se deduce que el modelo detecta bastante bien la innovación. Los resultados del modelo para datos test nunca vistos han sido positivos, ya que la interpretación de todas las métricas ha sido buena. A modo de conclusión, y como resumen de todas ellas, el modelo está proporcionando una precisión en torno al 66 %. Lo que implica que el modelo puede medir con cierto grado de fiabilidad si en un centro escolar se da de forma real innovación. Por tanto, se ha superado el objetivo principal de crear un modelo de ML, a partir de un conjunto de datos, que permita detectar la innovación educativa.

El segundo objetivo de este trabajo era diseñar un modelo que nos permita predecir el grado de innovación en una determinada institución educativa en función a unas variables objeto de estudio. Para alcanzar este objetivo se ha diseñado una web que se ofrece en abierto a cualquier usuario interesado en predecir el grado de innovación de una determinada institución educativa y que puede consultarse en el siguiente enlace: <https://658abb2e2582c934ce19e35--cute-pie-ff905a.netlify.app>. La construcción de esta aplicación predictora se basa en el modelo entrenado y validado que se ha presentado en este trabajo y constituye un elemento inédito en el panorama educativo de habla hispana.

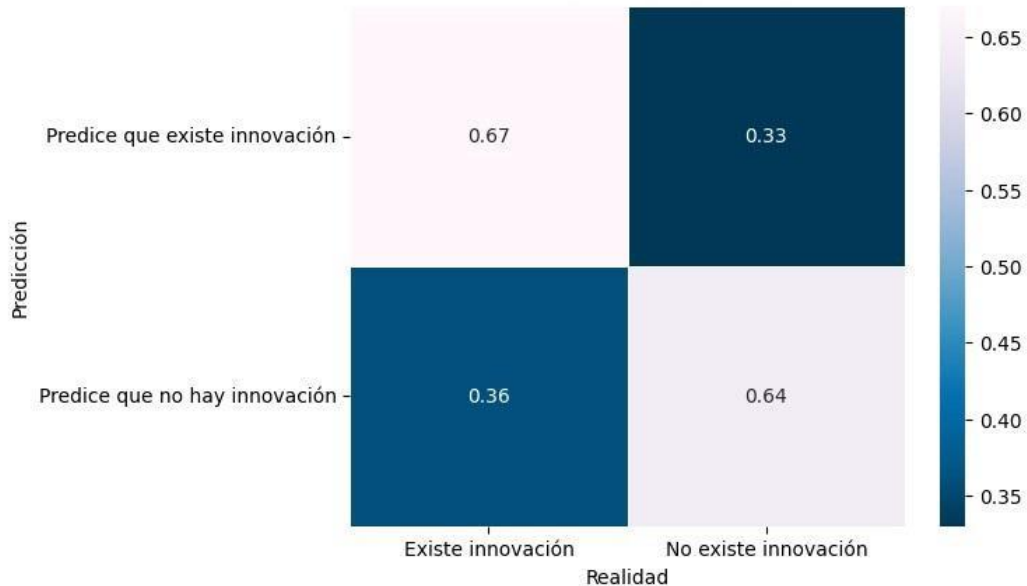


Fig. 4: Tabla de contingencia proporción por filas

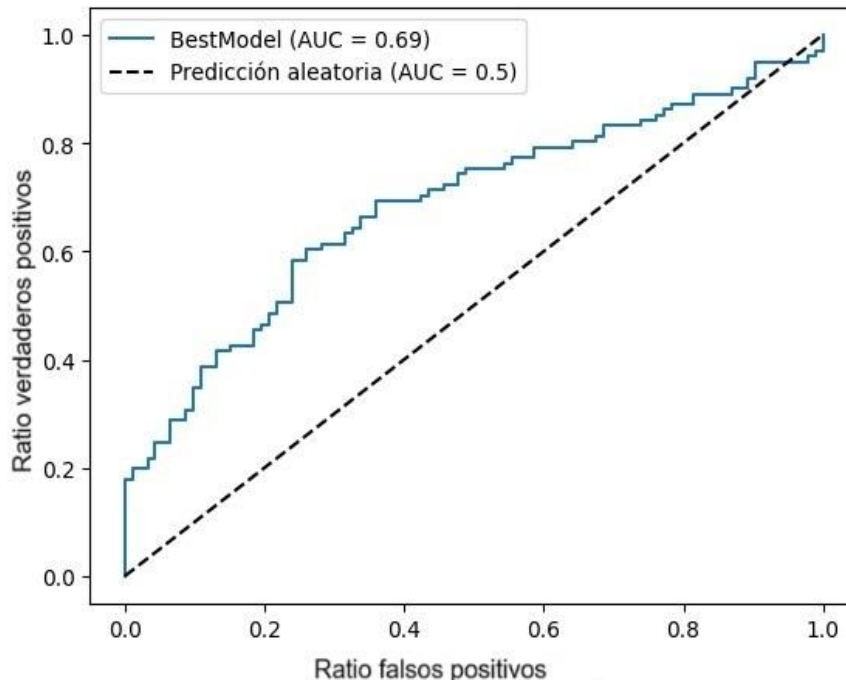


Fig. 5: Curva ROC

Identificación de Variables significativas

Para identificar las variables con mayor influencia en un modelo durante una evaluación, se utilizará la biblioteca SHAP (Shapley Additive exPlanations) en Python. Este método permitirá calcular las 10 variables más significativas y presentarlas visualmente mediante un diagrama de barras (Lundberg y Lee, 2017).

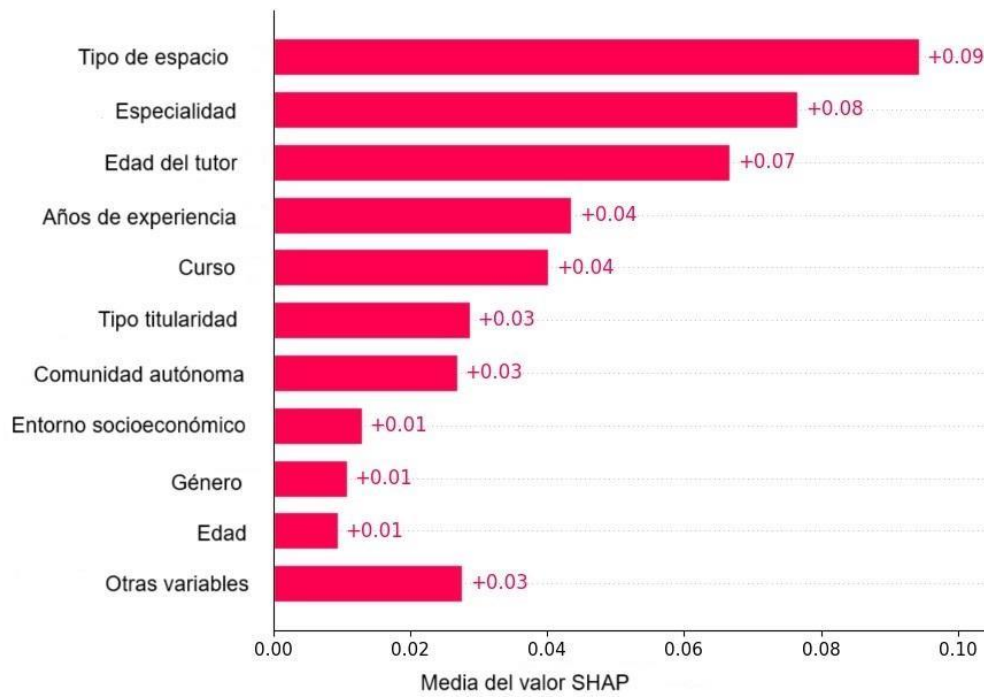


Fig. 6: Las 10 variables influyentes en el modelo

La Fig. 6 indica que las variables más influyentes son el tipo de espacio, la especialidad, el curso, el semestre, la edad del tutor y la formación. Por otro lado, es interesante observar las 10 variables más influyentes y su representación en un gráfico de enjambre de abejas dependiendo de su predicción. Las características también están ordenadas por su efecto en la predicción. Se puede ver cómo los valores más altos (rojo) y más bajos (azul) de las diferentes características afectarán al resultado. Como las variables son dicotómicas sólo puede tomar dos colores. Cada observación se representa con un punto de uno de los dos colores.

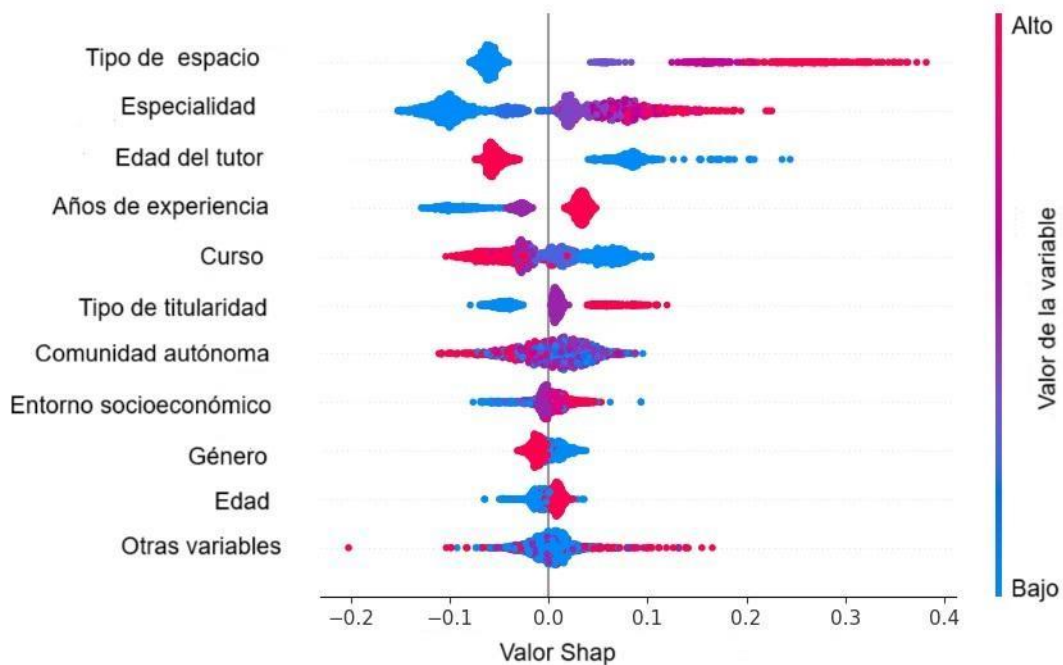


Fig. 7: Las 10 variables más influyentes del modelo en enjambre de abejas

En la Fig. 7 observamos que el eje horizontal representa el valor SHAP, mientras que el color del punto nos muestra si esa observación tiene un valor mayor o menor, en comparación con otras observaciones. Si el punto se sitúa a la izquierda no existe innovación y en caso de estar a la derecha sí. Por ejemplo, si la variable *Tipos espacios* tiene un valor bajo esto ayuda a que la predicción sea un 0, lo que significa que ayuda a que no haya innovación. Si es alto entonces ayuda a que la predicción sea un uno, es decir, que exista innovación educativa.

DISCUSIÓN FINAL

Destacamos que, a lo largo del presente trabajo, se ha proporcionado un modelo como resultado del uso de ML e IA. Este modelo ha sido debidamente entrenado en base a una serie de datos masivos proporcionados por el Observatorio de la Innovación Educativa. Tras el entrenamiento, hacemos constatar que nuestro modelo es capaz de medir el grado de implementación de propuestas metodológicas activas con una sensibilidad (esto es, la proporción de medir correctamente la innovación, sabiendo que de verdad se innova) del 67%. El código fuente asociado al modelo se encuentra en <https://github.com/FernandoCabezaSanchez/article-education-notebook.git>.

Adicionalmente, el modelo entrenado ha permitido discutir sobre la influencia de las variables a la hora de medir y predecir el uso de metodologías activas en una institución, lo que supone una temática y enfoque no abordado en otros trabajos sobre ML en educación como puede verse en la revisión bibliográfica realizada por Sanusi *et al.* (2023). Es interesante que gracias a esta técnica de análisis se haya visto que variables aparentemente no relacionadas directamente con el uso de metodologías activas tienen un impacto importante en su presencia en los centros educativos. La organización del espacio dentro de las aulas, la especialidad o asignatura que se imparte, el curso escolar en el que se trabaja, el semestre observado, la edad y la formación del profesor parecen ser variables importantes que deberían estudiarse mediante otras técnicas de análisis de datos para estudiar su significatividad y su posible influencia a la hora de desplegar políticas de innovación educativa en los centros educativos tal y como señalan Sghir *et al.* (2023). Este resultado de nuestro modelo y enfoque puede abrir nuevas técnicas de análisis de datos apoyadas en ML e IA que complementen los enfoques tradicionales basados en la estadística y es consistente con otros trabajos de ML en educación como los de Ong (2022) que también han servido para detectar variables no evidentes.

Un elemento para destacar de este trabajo es la creación de la página web donde se puede realizar una predicción de la capacidad de un centro educativo de innovar en metodologías activas en base a variables observables. Esta página supone un primer paso en la difusión y transferencia de la investigación educativa a la sociedad y permite que el modelo entrenado trascienda los límites de la investigación teórica y pueda ser empleado en la práctica y en la evaluación de la innovación educativa.

Un aspecto a tener en cuenta es que la innovación educativa puede ser entendida desde diferentes puntos de vista relacionados con la metodología, la organización, los recursos y la evaluación. En esta investigación solo se ha tenido en cuenta la primera de ellas, el uso de metodologías activas. Sin embargo, el estudio se podría ampliar de forma que se abordasen nuevos modelos que sirvieran de base para el análisis y la predicción de la innovación en la organización de los centros educativos, en el tipo de recursos que se emplean en las aulas o en el enfoque y en el alcance de la evaluación que se lleva a cabo. Por otro lado, se podría realizar un estudio que recogiese la innovación total desde los cuatro puntos a la vez. Todos esos estudios completarían éste y podrían enriquecer el modelo de previsión web creado.

Por otro lado, cabe preguntarse si un 66 % es un porcentaje admisible para un modelo de redes neuronales, es cierto que para muchos conjuntos de datos ese porcentaje de precisión parece insuficiente. Sin embargo, el conjunto de datos, a pesar de ser extenso, puede resultar pequeño para modelos de aprendizaje automático.

Para solventar las limitaciones desde cualquiera de los puntos de vista anteriores se debe aumentar el número de observaciones y añadir más variables que se puedan obtener con facilidad, por ejemplo, carrera universitaria del tutor de prácticas en el centro. Con la obtención de nuevos datos y el aumento de estos se podrá reentrenar a la red neuronal para mejorar la precisión. Adicionalmente mejorando las encuestas se puede obtener características fundamentales que aumenten la precisión. Por otro lado, es una puerta de entrada a muchos otros estudios de una índole similar, por ejemplo, estudiar el rendimiento académico de los alumnos mirando su histórico de notas y ciertas características contextuales como se hizo en Duzhin y Gustafsson (2018), estudiar la retención académica como se investigó en Musso *et al.* (2020) y da respuesta a la necesidad señalada por Urbina y De la Calleja (2017) de ampliar el uso del ML al ámbito educativo.

A pesar de todos estos inconvenientes el modelo funciona con un porcentaje razonable. En definitiva, esta investigación ha sido innovadora y puede mejorar la educación haciendo que los estudiantes elijan bien el centro en el que realizar las prácticas o dando herramientas a las familias para elegir centros educativos que estén en una mejor posición de innovar en metodologías activas. Finalmente, se ha superado el objetivo principal de crear un modelo de ML, a partir de un conjunto de datos, que permita detectar la innovación educativa en metodologías activas con una precisión cercana al 66% y se ha creado una web basada en el modelo que nos permite predecir el grado de innovación en una determinada institución educativa en función de unas variables observables y públicas.

CONCLUSIONES

De acuerdo al trabajo presentado y a los resultados obtenidos, se pueden plantear las siguientes conclusiones principales:

1. Validación del Modelo: El modelo de aprendizaje automático desarrollado ha demostrado ser eficaz para predecir el grado de innovación en centros educativos, alcanzando una precisión del 66%, lo cual indica su potencial para ser aplicado en entornos reales para evaluar y fomentar prácticas innovadoras en la educación.
2. Impacto de las Variables: La investigación ha identificado variables críticas que influyen significativamente en la capacidad de innovación de los centros, como el tipo de espacio, la especialidad impartida y la formación del profesorado. Este hallazgo subraya la importancia de considerar estos factores al implementar y evaluar metodologías educativas activas.
3. Herramienta de Predicción: La creación de una herramienta web basada en el modelo permite a usuarios externos estimar el nivel de innovación de instituciones educativas. Esto facilita una valoración accesible y rápida que puede ser utilizada por administradores y educadores para planificar mejoras y cambios estratégicos.
4. Ampliación del Modelo: Aunque el modelo actual es prometedor, los resultados sugieren que podría mejorarse mediante la integración de más datos y la exploración de nuevas variables. Esto podría aumentar la precisión del modelo y proporcionar insights más detallados sobre los factores que impulsan la innovación educativa.
5. Implicaciones Futuras: Este estudio abre la puerta a futuras investigaciones que podrían expandir el modelo a otros aspectos de la innovación educativa. En concreto sería muy interesante tratar de abordar más a fondo la precisión del modelo y su aceptabilidad en el contexto educativo y comparar el porcentaje de precisión del modelo con otros modelos similares en el campo. A partir de estos puntos podría mejorarse en el futuro y discutir la validez y la capacidad de generalización del modelo, explorando sus limitaciones y posibles sesgos.

Estas conclusiones resaltan la relevancia del estudio y su aplicación potencial en la mejora continua de prácticas educativas innovadoras y sientan las bases para múltiples líneas de investigación futuras.

REFERENCIAS

- Alyahyan, E., y Düşteğör, D., Predicting academic success in higher education: literature review and best practices, <https://doi.org/10.1186/s41239-020-0177-7>, Int J Educ Technol High Educ, 17, 3 (2020).
- Assi, K.J., Shafiullah, M., Nahiduzzaman, K.M., y Mansoor, U., Travel-To-School Mode Choice Modelling Employing Artificial Intelligence Techniques: A Comparative Study, <https://doi.org/10.3390/su11164484>, Sustainability, 11(16), 1-13 (2019).
- BOE, Real Decreto 217/2022, de 29 de marzo, por el que se establece la ordenación y las enseñanzas mínimas de la Educación Secundaria Obligatoria, <https://www.boe.es/eli/es/rd/2022/03/29/217/con>, Boletín Oficial del Estado, España (2022)
- BOE, Real Decreto 243/2022, de 5 de abril, por el que se establecen la ordenación y las enseñanzas mínimas del Bachillerato, <https://www.boe.es/eli/es/rd/2022/04/05/243/con>, Boletín Oficial del Estado, España (2022)
- Cabero, J., y Barroso, J., La utilización del juicio de experto para la evaluación de TIC: el coeficiente de competencia experta, <https://recyt.fecyt.es/index.php/BORDON/article/view/brp.2013.65202>, Bordón, 65(2), 25-38 (2013)
- Contreras, L., Fuentes, H., y Rodríguez, J., Predicción del rendimiento académico como indicador de éxito/fracaso de los estudiantes de ingeniería, mediante aprendizaje automático, <http://dx.doi.org/10.4067/S0718-50062020000500233>, Formación Universitaria, 13(5), 233-246 (2020)
- Díaz, J.L., Sánchez, A., y Roa, J., Estado de uso de metodologías activas en las aulas de matemáticas secundarias, Journal of Research in Mathematics Education, 12(3), 229-245 (2023).
- Duzhin, F., y Gustafsson, A., Machine learning-based app for self-evaluation of teacher-specific instructional style and tools, <https://doi.org/10.3390/educsci8010007>, Education Sciences, 8(1), 1-15 (2018).
- Hale, C., Data Transformation: Techniques and best practices, Journal of Data Management, 34(2), 117-132 (2021).
- Little, R. J., y Rubin, D. B., Statistical analysis with missing data, 3ª edición, Wiley (2020)
- Lundberg, S. M., y Lee, S. I., A Unified Approach to Interpreting Model Predictions, <https://doi.org/10.48550/arXiv.1705.07874>, arXiv:1705.07874 (2017).
- MEFP, Ministerio de Educación y Formación Profesional, Las cifras de la educación en España, curso 2019-2020 (2022 ed.), <https://sede.educacion.gob.es/publiventa/d/26071/19/0> (2021)

Mosquera, R., Castrillón, D., y Parra, L., Predicción de Riesgos Psicosociales en Docentes de Colegios Públicos Colombianos utilizando Técnicas de Inteligencia Artificial, <http://dx.doi.org/10.4067/S0718-07642018000400267>, Inf. Tecnol., 29(4), 267-280 (2018)

Musso, M.F., Hernández, C.F.R., y Cascallar, E.C., Predicting key educational outcomes in academic trajectories: a machine-learning approach, <https://doi.org/10.1007/s10734-020-00520-7>, High Educ, 80, 875–894 (2020)

Ong, A.K.S., A Machine Learning Ensemble Approach for Predicting Factors Affecting STEM Students' Future Intention to Enroll in Chemistry-Related Courses, <https://doi.org/10.3390/su142316041>, Sustainability, 14(23), 16041 (2022)

Raschka, S., Patterson, J., y Nolet, C., Machine learning in python: Main developments and technology trends in data science, machine learning, and artificial intelligence, <https://doi.org/10.3390/info11040193>, Information, 11(4), 1-44 (2020)

Sanusi, I.T., Oyelere, S.S., Vartiainen, H., Suhonen, J. y Tukiainen, M., A systematic review of teaching and learning machine learning in K-12 education, <https://doi.org/10.1007/s10639-022-11416-7>, Educ Inf Technol, 28, 5967–5997 (2023)

Sarker, I. H., Machine learning: Algorithms, real-world applications and research directions, SN Computer Science, 2(3), 121 (2021)

Sghir, N., Adadi, A., y Lahmer, M., Recent advances in Predictive Learning Analytics: A decade systematic review (2012–2022), <https://doi.org/10.1007/s10639-022-11536-0>, Educ Inf Technol, 28, 8299–8333 (2023)

Tharwat, A., Classification assessment methods, <https://doi.org/10.1016/j.aci.2018.08.003>, Applied Computing and Informatics, 17(1), 168-192 (2021)

Urbina, A. B., y De la Calleja, J., Brief review of educational applications using data mining and machine learning, <https://doi.org/10.24320/redie.2017.19.4.1305>, Revista Electrónica de Investigación Educativa, 19(4), 84-96 (2017)

Witten, I. H., Frank, E., Hall, M. A., y Pal, C. J., Data Mining: Practical Machine Learning Tools and Techniques, 4ª edición, Morgan Kaufmann (2016)