

Evaluación económica de intervenciones educativas en la LOMLOE*: propuestas de mejora con inteligencia artificial

An economic evaluation of educational interventions in the LOMLOE: Proposals for improvement with Artificial Intelligence

Dra. María Teresa BALLESTAR. Profesora Asociada. Universidad Rey Juan Carlos (teresa.ballestar@urjc.es).

Dr. Jorge SAINZ. Catedrático. Universidad Rey Juan Carlos (jorge.sainz@urjc.es).

Dr. Ismael SANZ. Profesor Titular. Universidad Rey Juan Carlos (ismael.sanz@urjc.es).

Resumen:

El objetivo de esta investigación es demostrar la necesidad de evaluar económicamente la LOMLOE, especialmente tras la inversión de los fondos *EU Next Generation* que abren nuevas oportunidades de las que carecía la ley en su redacción inicial. Las Administraciones públicas tienen el reto de emplear esa inversión adicional de forma eficiente.

Nuestro análisis demuestra que los modelos de inteligencia artificial pueden predecir si los programas de apoyo educativo ayudarán a incrementar la probabilidad de que estudiantes rezagados superen 4.º de la ESO (Educación Secundaria Obligatoria). De esta forma, se puede calcular el retorno social de los pro-

gramas de apoyo educativo y contribuir a su diseño *ex-ante* para lograr que los alumnos tengan mayores tasas de éxito.

Para completar los modelos ya utilizados por Administraciones públicas, empleamos modelos de Machine Learning (ML) robustos como árboles de decisión CHAID y redes neuronales artificiales para analizar las características de los grupos de estudiantes y la intervención en la que han formado parte. Las conclusiones permiten mejorar los programas de refuerzo educativo de los próximos años para apoyar a los alumnos con menos posibilidades de éxito académico.

Descriptores: análisis de políticas públicas, *Machine Learning*, eficiencia educativa, LOMLOE.

* Ley Orgánica 3/2020, de 29 de diciembre de 2020, por la que se modifica la Ley Orgánica 2/2006, de 3 de mayo, de Educación. Fecha de recepción de la versión definitiva de este artículo: 13-12-2021.

Cómo citar este artículo: Ballestar, M. T., Sainz, J. y Sanz, I. (2022). Evaluación económica de intervenciones educativas en la LOMLOE: propuestas de mejora con inteligencia artificial [*An economic evaluation of educational interventions in the LOMLOE: Proposals for improvement with Artificial Intelligence*]. *Revista Española de Pedagogía*, 80 (281), 133-154. <https://doi.org/10.22550/REP80-1-2022-09>
<https://revistadepedagogia.org/>

Abstract:

This research aims to demonstrate the need for an economic evaluation of the Organic Law 3/2020, of 29 December, which amends Organic Law 2/2006, of 3 May, on Education (LOMLOE), especially after the investment of EU Next Generation funds that open new opportunities that were lacking in the initial drafting of the law. The challenge for Public Administrations is to use this additional investment efficiently.

Our analysis shows that artificial intelligence models can predict whether educational support programmes will help increase the likelihood that students who lag behind will pass the 4th grade of ESO. In this way, we can calculate the social return of these

programmes and contribute to their *ex-ante* design to achieve higher success rates for students.

To complement the models already used by public administrations, we use robust Machine Learning (ML) models such as CHAID decision trees and artificial neural networks to analyse the characteristics of the groups of students and the intervention they have been part of. The conclusions allow us to improve educational reinforcement programmes in the coming years to support students with lower chances of academic success.

Keywords: Public policy analysis, Machine Learning, educational efficiency, LOMLOE.

1. Introducción

Uno de los principios en la gestión educativa por parte de organismos internacionales como la OCDE (Organización para la Cooperación y el Desarrollo Económicos) es la evaluación de políticas educativas tanto globales como de intervenciones concretas (Golden, 2020). El resultado del análisis lleva a una mayor eficacia e impacto, tanto a nivel social como individual, de las actuaciones educativas, lo que revierte en el alumnado y en una dinámica positiva para el sistema en general (OECD, 2018). Este principio informativo aparece en casi todas las legislaciones europeas y, en el caso español, se encuentra recogido en Ley Orgánica 2/2006, de 3 de mayo (LOE) que señala en su Artículo 2 bis que el «El funcionamiento del Sistema Educativo Espa-

ñol se rige por los principios de [...] eficiencia en la asignación de recursos públicos, transparencia y rendición de cuentas».

Nuestro objetivo es demostrar que la evaluación económica de las políticas educativas sirve para optimizar y priorizar los objetivos legislativos establecidos por la LOMLOE. Este hecho demostraría que la ley se enfrenta a una laguna en su construcción que afecta a su evaluación de la gestión, evaluación que si es un requerimiento para la justificación de los fondos *EU Next Generation* para educación que solo en 2021 representan en los presupuestos generales del estado 1852.5 millones de euros.

Va a ser esta inversión europea la que permitirá implantar y evaluar la ley, ya

que en su Memoria de Análisis e Impacto Normativo (MAIN), que recoge los aspectos técnicos de la norma, se señala que «[...] el proyecto no supone incremento o disminución del gasto público». Como contrapartida la Comisión exige que se evalúe si la población adquiere efectivamente las competencias para competir a nivel global (Crescenzi et al., 2021; Porte y Jensen, 2021).

La Comisión Europea y UNESCO reconocen la importancia de este tipo de actuaciones tanto en la elección de reformas como en su posterior análisis, y la necesidad de que haya datos disponibles sobre intervenciones educativas para su posterior análisis, pese a las reticencias de algunas administraciones (European Commission, 2010; Yusuf, 2007). Aunque el planteamiento preferido para este análisis es el enfoque experimental, en la realidad educativa existen problemas para llevarlo a cabo por diversos motivos: como el coste económico, la dificultad de establecer grupos de control aleatorios, los problemas de gestión, ansiedad, etc. (Golden, 2020; Slavin, 2016).

Para solventar esta propuesta se están utilizando un arsenal de nuevas técnicas que están en pleno desarrollo y que pueden servir como propuesta, entre las que se encuentra la inteligencia artificial (IA) (Ballestar et al., 2019; Chassignol et al., 2018; Chatterjee y Bhattacharjee, 2020). Siguiendo la estrategia de triangulación de Ballestar et al. (2020), evaluamos el uso de distintas estrategias de IA para la evaluación del retorno de intervenciones educativas. Para ello, profundizaremos en la intervención realizada por la Consejería de Educación de la Junta de Comunidades

de Castilla y León denominada *Programa para la Mejora del Éxito Educativo*, que analizaremos en su edición de 2020, tras el cierre de los colegios por la pandemia. La Junta realizó una convocatoria de ayudas económicas para que los centros educativos pudieran ofrecer aulas de apoyo durante el mes de julio a los alumnos de sexto de primaria, cuarto de ESO y segundo de Bachillerato en las áreas de lengua castellana y literatura, matemáticas o inglés con dificultades educativas. El programa ha contribuido a lo largo de los años a que un número importante de alumnos de la comunidad promocionen de curso. De forma general para todos los cursos que disfrutaban del programa, se observó en el curso académico 2019-20 una mejora de la promoción del 5 % en los estudiantes que participaron en el programa, en comparación con los que estando en condiciones de participar no lo hicieron.

2. La intervención

Para diseñar cualquier intervención educativa Slavin (2016) y Golden (2020) reconocen como fundamental que esté basada en evaluaciones científicas. Probablemente, las clases de apoyo que componen el *Programa para la Mejora del Éxito Educativo* en Castilla y León sea una de las más fundamentadas científicamente en los últimos años. Las investigaciones realizadas por la *Education Endowment Foundation*¹ demuestran que las tutorías en pequeños grupos se encuentran entre las medidas para las que existe evidencia empírica de mayor efectividad a nivel internacional. Como tal han demostrado ser un buen complemento para un sistema educativo que está diseña-

do para hacer pasar a un gran número de alumnos de curso a curso, pero que no funciona para todos los estudiantes.

Kraft (2015) y Burgess (2020) analizan los factores que contribuyen a su éxito: la impartición por personal seleccionado y formado, coordinado con sus docentes habituales y en grupos reducidos. El impacto es relevante: Nickow et al., (2020) muestran, a partir del estudio de 96 artículos de experimentos aleatorios, que el efecto de las tutorías en pequeños grupos es importante y significativo (37 % de la desviación estándar).

Este tipo de intervenciones es una de las pocas que vienen recogidas en nuestra legislación. La LOMLOE en su art. 4.4 señala que uno de sus objetivos es facilitar «[...] el acceso a los apoyos que el alumnado requiera» y el incremento de estas medidas a lo largo de todo el proceso formativo de forma individualizada (art. 20 bis) «[...] para evitar la repetición escolar, particularmente en entornos socialmente desfavorecidos». De hecho, la norma deja una redacción muy clara sobre sus objetivos a la repetición, por ejemplo, en la nueva redacción del art. 28: «La permanencia en el mismo curso [de secundaria] se considerará una medida de carácter excepcional y se tomará tras haber agotado las medidas ordinarias de refuerzo y apoyo para solventar las dificultades de aprendizaje».

Estas medidas de apoyo están presentes en la legislación educativa a lo largo de este siglo. Así la *non nata* Ley Orgánica 10/2002, de 23 de diciembre, de Calidad de la Educación (LOCE) recogía en su art. 2 el derecho de los estudiantes a «[...] recibir las ayudas

y los apoyos precisos para compensar las carencias y desventajas de tipo personal, familiar, económico, social y cultural, especialmente en el caso de presentar necesidades educativas especiales, que impidan o dificulten el acceso y la permanencia en el sistema educativo». Las distintas normas desde entonces han recogido preceptos similares de protección en línea con las recomendaciones de los organismos especializados (Gouédard et al., 2020; Pont y Montt, 2014; Schleicher, 2020) como la literatura sobre el retorno educativo (Brunello y Paola, 2014; de la Fuente y Jimeno, 2009; Doncel et al., 2014) que establecen la importancia de la adquisición de competencias, hecho que pasa a ser fundamental con la irrupción de las nuevas tecnologías (Ballestar et al., 2020, 2022; Goos et al., 2009; Gregory et al., 2019).

De acuerdo con el art. 9 de la LOMLOE corresponde a las comunidades autónomas (CCAA) llevar a cabo las actuaciones dirigidas a la reducción y prevención del fracaso escolar y del abandono escolar temprano a través de los programas de cooperación territorial. La Memoria de Análisis e Impacto Normativo de la Ley dotó en la redacción original de la ley, previa a la pandemia, con 45 millones de euros por año para este tipo de programas durante el periodo 2020-2023.

Las CCAA ya venían destinando una importante inversión a este tipo de proyectos. La Consejería de Educación de Castilla y León venía desarrollando desde el curso 2007-2008, entre otros, el *Programa para la Mejora del Éxito Educativo*. Sus objetivos son, en línea con la LOE, LOMCE (Ley Orgánica para la Mejora de la Calidad

Educativa, 2013) y LOMLOE «[...] contribuir a la mejora del modelo educativo en la Comunidad de Castilla y León y facilitar el éxito y la continuidad del alumnado en el sistema educativo» garantizando «la progresión didáctica del alumnado y su promoción en el sistema educativo, reduciendo el abandono escolar temprano, fomentando su eficaz inserción en el mercado laboral y optimizando al mismo tiempo el clima de convivencia en los centros docentes».

Son los centros los que solicitan estas medidas para favorecer el éxito educativo, la integración del alumnado y la involucreción de las familias para aumentar la tasa de promoción y titulación del alumnado, con especial atención al alumnado más vulnerable de 4.º de ESO, 6.º curso de Educación Primaria y 2.º de Bachillerato de centros sostenidos con fondos públicos. El programa se imparte en 93 centros en lengua, matemáticas e inglés de cara a reforzar el fin de ciclo y con la participación de estudiantes que asisten voluntariamente, autorizados y apoyados por su familia. Los requisitos de acceso son presentar una necesidad específica de apoyo educativo o estar en situación de vulnerabilidad socioeducativa y es impartida por docentes especializados.

La primera evaluación realizada por la junta es prometedora: para el curso 2019-2020 los alumnos que participan en el programa aumentan su probabilidad de aprobar lengua en un 24.6 %, inglés en un 13.8 % y matemáticas en un 8.4 % frente a los no inscritos, generando un alto nivel de satisfacción tanto en docentes (85.7 %) como familias (88.2 %).

Este primer análisis nos sirve como base para realizar nuestra investigación sobre la eficiencia económica de las intervenciones recogidas en la LOMLOE. Para ello, planteamos clasificar en grupos a los estudiantes que participan en el *Programa para la Mejora del Éxito Educativo* en Castilla y León en 4.º de la ESO y predecir la probabilidad de superar dicho curso escolar para cada uno de los grupos de estudiantes. Esto permitirá realizar pronósticos individuales para estudiantes en base a sus características, detalles del programa especial de apoyo realizado y su desempeño en el programa. Esto permite estimar la probabilidad de éxito del programa y calcular su retorno social (ROI) a través de un modelo de triangulación que emplea diferentes metodologías de ML como árboles de decisión CHAID y redes neuronales artificiales para evaluar políticas educativas.

3. Análisis empírico

La información empleada para el análisis contiene los datos disponibles de los alumnos de 4.º de la ESO participantes en el *Programa para la Mejora del Éxito Educativo* en el curso 2019-2020. Se dispone de 1739 registros que corresponden con los estudiantes que participaron en alguno de los tres programas implementados en la comunidad autónoma. De ellos, el 47.27 % son niñas y el 52.73 % son niños. De estos estudiantes, el 47.15 % (820 estudiantes) han repetido ya algún curso.

Los estudiantes pueden participar en tres tipos de programas. El programa C2 consiste en un acompañamiento del estudiante durante todo el año académico

y acumula el 76.37 % de los estudiantes (1328). El programa C3 consiste en la asistencia de clases extraordinarias durante verano en el mes de julio e incluye al 17.02 % de los estudiantes (296). Finalmente, el programa C2C3, una combinación de los dos anteriores, representa el 6.61 % de los estudiantes (115). Dichas actuaciones se imparten en las nueve provincias de Castilla y León, concentrando el 66.36 % de los estudiantes en capitales de provincia y el 33.64 % en otras localidades, con una ratio de éxito en la finalización de los estudios de 4.º de la ESO del 85.34 % para aquellos que terminaron el programa.

Una de las innovaciones de nuestra investigación es la aplicación de una metodología de triangulación que consiste en el desarrollo de más de un método cuantitativo, aplicando diferentes enfoques, con el objetivo de enriquecer los resultados, así como, confirmar por partida doble los resultados obtenidos (Ballestar et al., 2020). En la primera fase, se desarrolla un modelo de Machine Learning (ML) basado en árboles de decisión CHAID (Chi-square Automatic Interaction Detector) que determine cuáles son las variables relevantes a la hora de diseñar y predecir la tasa de éxito esperada que tendrán los participantes en el programa. El objetivo es identificar los factores a tener en cuenta a la hora de diseñar una intervención de estas características, así como calcular su tasa de éxito y, en consecuencia, el retorno social de la inversión realizada.

En la segunda fase, el primer modelo se evalúa mediante la realización de un

modelo completamente nuevo que utiliza una metodología ML basada en redes neuronales artificiales perceptrón multicapa con programación hacia atrás (ANN-MPL) que debe validar los resultados obtenidos anteriormente.

3.1. Definición de las variables de los modelos de ML

Los árboles de decisión CHAID y las redes neuronales artificiales perceptrón multicapa (ANN-MLP) son métodos de aprendizaje automático supervisados. Este tipo de modelos describen y explican las relaciones subyacentes existentes entre las diferentes variables de entrada con la finalidad de predecir el valor de la variable objetivo, a través de un proceso de entrenamiento usando una muestra de datos que contiene tanto los valores para las variables de entrada como de salida (Maimon y Rokach, 2005). Como métodos supervisados, la precisión de sus clasificaciones y predicciones estará condicionada a la calidad de la muestra disponible para el entrenamiento, validación y test de los modelos (Aad et al., 2012; Ballestar et al., 2018; Li y Eastman, 2006).

Se ha realizado un proceso de minería de datos sobre la base de origen con el objetivo de crear una única tabla para el desarrollo de los modelos de ML. Esta tabla contiene 1739 registros de estudiantes y 21 variables que capturan la información para la caracterización del estudiante, así como su desempeño y resultados tanto en el programa que ha participado como en el curso de 4.º de la ESO.

Mediante la realización de análisis descriptivos y causales iterativos, 5 de estas variables fueron identificadas como estadísticamente significativas y relevantes para el análisis empírico y desarrollo de los dos modelos de ML. De estas variables, 4 actúan como variables de entrada de información en los modelos y 1 como variable de salida u objetivo a determinar o predecir (Tabla 1). Por lo tanto, ambos modelos

utilizan las mismas 5 variables y son entrenados con la misma base de datos.

La variable objetivo o de salida, el porcentaje de estudiantes que no ha superado el curso de 4.º de la ESO (0 en la variable *finish_studies*), es 14.66 %, mientras que los que sí superan el curso que representan el 85.34 % (valor 1 de la variable *finish_studies*), dejando una muestra desbalanceada.

TABLA 1. Descripción de las variables de los modelos de ML: árbol de decisión CHAID y red neuronal artificial perceptrón multicapa (ANN-MLP).

Variables de entrada	Descripción
years_repetition	Variable numérica discreta. Número de cursos que ha repetido el estudiante. En el caso de que el estudiante no haya repetido ningún curso el valor será 0.
student_repetition	Variable booleana. Valor 1 si el estudiante ha repetido algún curso; Valor 0 si no ha repetido ningún curso.
kind_program	Variable categórica que indica cuál de los tres programas académicos ha realizado el estudiante: C2; C3; C2C3.
finish_program	Variable booleana. Valor 1 si el estudiante ha finalizado el programa académico de apoyo; Valor 0 si no lo ha llegado a finalizar.
Variable de salida	Descripción
finish_studies	Variable booleana. Valor 1 si el estudiante ha superado el curso 4º de la ESO; Valor 0 si no ha llegado a superar el curso.

Fuente: Elaboración propia.

3.2. Primera fase del análisis empírico: modelo predictivo del éxito del programa especial de apoyo a la educación

En esta fase se desarrolla un modelo consistente en un árbol de decisión CHAID o *chi-squared automatic interaction detection* (Kass, 1980), cuyo objetivo es doble: clasificación y predicción. Este modelo clasificará a cada uno de los estudiantes en grupos homogéneos en base a las variables explicativas de entrada o predictores

que interactúan significativamente con la variable dependiente de salida u objetivo (Akin et al., 2017; Khosravi et al., 2019). El algoritmo identifica los estudiantes que probablemente pertenezcan a un grupo en particular y, al mismo tiempo, identifica las reglas que utilizará en predicciones de futuros casos de estudiantes que no hayan participado. En nuestro caso estas agrupaciones están constituidas en función de las características del estudiante, del programa

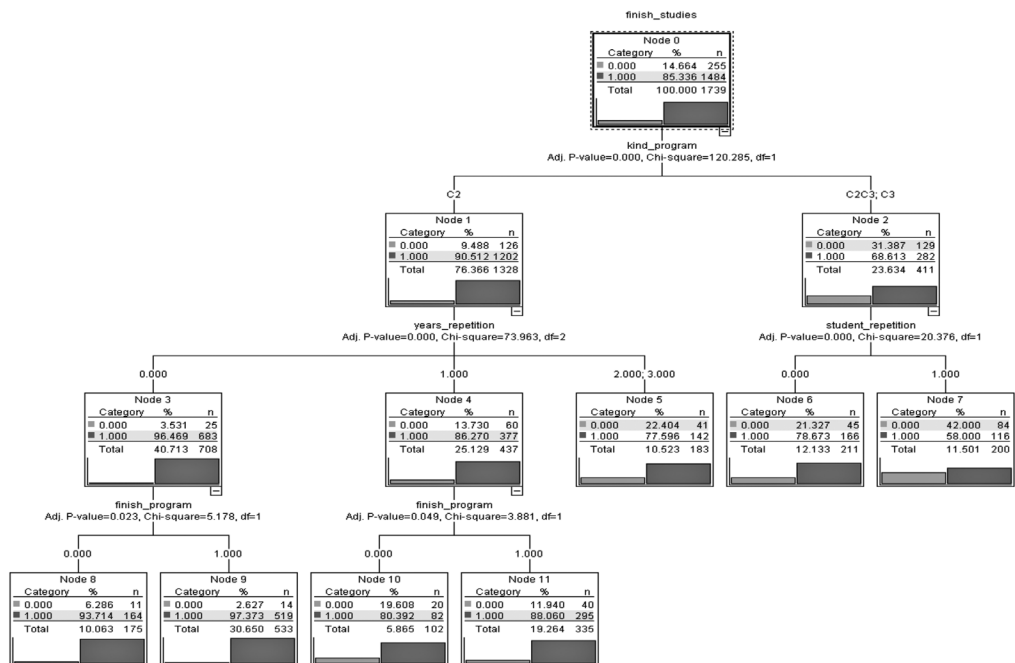
en el que participa y su desempeño, y posibilitan cumplir el segundo objetivo de determinar y predecir el porcentaje de aprobados en 4.º de la ESO tanto para cada uno de los grupos como individualmente (Ramasmami y Bhaskaran, 2010). Finalmente, la predicción de este porcentaje de éxito permite calcular como reforzar cada uno de los programas por tipología de estudiante.

La elección de árboles de decisión CHAID frente a otros métodos, como los propuestos para variables educativas por, por ejemplo, Cruz-Jesus et al., (2020) es que tienen la capacidad de generar árboles no binarios. Además, también admiten variables de entrada tipo numérico continuas o categóricas. Al ser no paramétrico puede manejar tanto relaciones lineales

como no lineales entre las variables explicativas y la variable de salida y también puede gestionar de forma muy eficiente grandes volúmenes de datos, incluso en tiempo real (Chassignol et al., 2018; Khosravi et al., 2019).

Para el desarrollo y entrenamiento del modelo se ha utilizado una metodología de entrenamiento, comprobación y validación (*training, testing, validación* [TTV]), que consiste en entrenar el modelo con el 70 % (1203 registros) de la muestra y realizar la comprobación y su validación con el restante 20 % (367 registros) y 10 % (169 registros) de la muestra respectivamente (Ballestar et al., 2019). El Gráfico 1 muestra el diagrama de árbol de decisión CHAID obtenido para el modelo de éxito del programa.

GRÁFICO 1. Diagrama de árbol de decisión CHAID para el modelo de éxito de los programas de apoyo a estudiantes para la superación del curso 4º. ESO.



Fuente: Elaboración propia.

3.2.1. Primera fase del análisis empírico: Evaluación del modelo.

Los indicadores más relevantes de CHAID son la precisión en la clasificación, sensibilidad, especificidad, el área debajo de la curva ROC y el coeficiente de GINI. Estos indicadores de precisión y la matriz de confusión, que contiene el porcentaje de casos clasificados tanto correcta como incorrectamente para los dos posibles valores de la variable dependiente, están disponibles en la Tabla 2. Los indicadores han sido calculados tanto para el total de la muestra, como para cada una de las submuestras de entrenamiento, comprobación y validación del modelo. Por lo tanto, observamos que el modelo tiene una precisión en la clasificación de estudiantes y predicción de su porcentaje de estudiantes aprobados del 70.73 % (error del 29.27 %). Este porcentaje de estudiantes clasificados correctamente es muy similar entre todas las submuestras de entrenamiento, comprobación y validación, por lo que podemos afirmar que el modelo no ha sido sobreentrenado.

El porcentaje de positivos verdaderos, denominado *sensibilidad*, es del 71.43 %. Este valor explica el porcentaje de estudiantes que superan el curso y que han sido correctamente clasificados por el árbol de decisión CHAID en base a las características del estudiante, el programa de apoyo realizado y su resultado en dicho programa. El porcentaje de negativos verdaderos, también denominado *especificidad*, es del 66.67 %. Este valor es el porcentaje de estudiantes que no superan 4.º de la ESO y que han sido correctamente clasificado en base a las mismas variables de entrada. Los valores complementarios son el porcentaje de falsos positivos, 33.33 %, corresponde con el porcentaje de estudiantes que no habiendo su-

perado 4º de la ESO fueron clasificados por el modelo como que si lo habían conseguido. Finalmente, el porcentaje de falsos negativos es del 28.57 %. Este valor corresponde al porcentaje de estudiantes que habiendo superado el curso fueron clasificados por el modelo como que no lo habían superado.

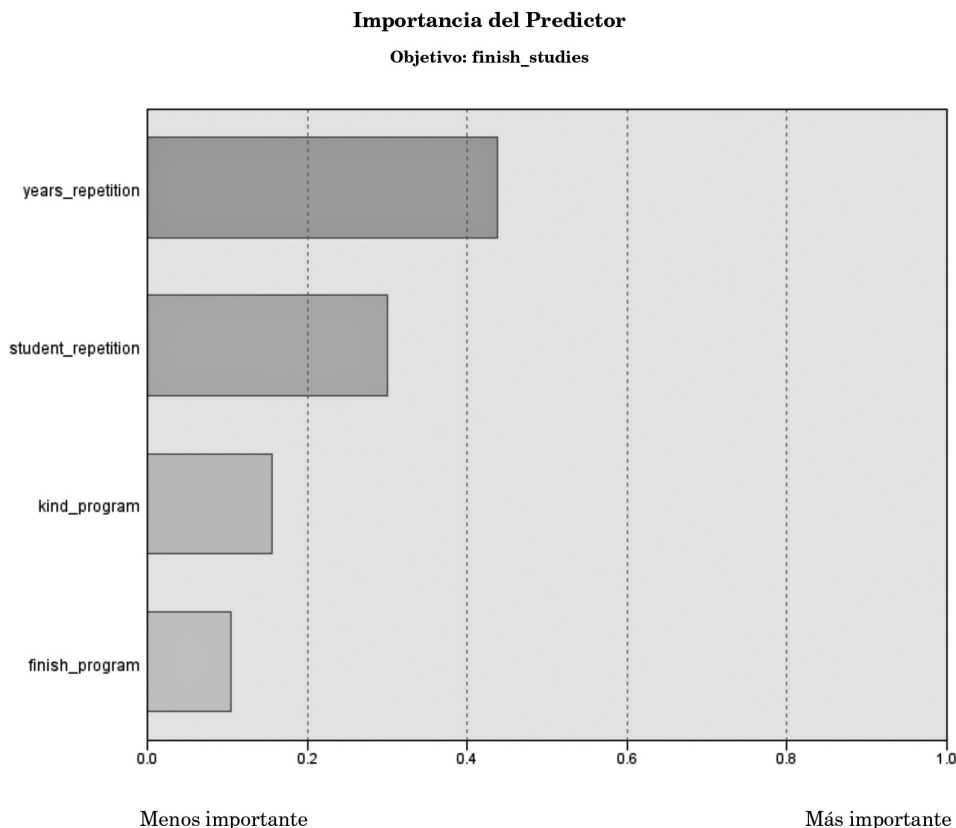
En esta investigación preferimos utilizar como principal medida de precisión del árbol de decisión CHAID el indicador de área debajo de la curva (AUC) ROC, ya que resulta más robusto que el indicador de precisión en la clasificación cuando se trabaja con muestras no balanceadas, como es nuestro caso (Tabla 2) (Dželihodžić y Jonko, 2016; Yin et al., 2013). El área debajo de la curva (AUC) ROC para el total de la muestra tiene un valor de 0.762, similar a las de las submuestras de entrenamiento, comprobación y validación del modelo (0.763, 0.760 y 0.755 respectivamente). Los valores de AUC pueden oscilar entre 0.5, que es el peor valor posible e implicaría que el modelo hace clasificaciones al azar y 1, que es el mejor valor y significaría que el modelo hace clasificaciones perfectas. Por lo tanto, se puede concluir que la calidad de este modelo de árbol de decisión CHAID es bueno (Hosmer Jr. et al., 2013). De forma complementaria, también se ha calculado el coeficiente de GINI, que se puede utilizar alternativamente al AUC, ya que están estrechamente relacionados. El coeficiente de GINI representa dos veces el área comprendida entre la curva ROC y la diagonal y oscila entre los valores 0 y 1. En nuestra investigación el coeficiente de GINI para el total de la muestra tiene un valor de 0.523, similar a la de las submuestras de entrenamiento, comprobación y validación del modelo (0.527, 0.519 y 0.510 respectivamente).

TABLA 2. Precisión del modelo y matriz de confusión del árbol de decisión CHAID.

Precisión del Modelo		Matriz de Confusión					
		Predicción					
		Tamaño muestra					
		Muestra			Porcentaje		
Muestra	Tamaño Muestra	Porcentaje Correcto	AUC	GINI	Observado	0	1
Total Muestra	1739	70.73 %	0.762	0.523	0	170	85
					1	424	1060
						28.57 %	71.43 %
Submuestra							
Entrenamiento	1203	70.91 %	0.763	0.527	Entrenamiento	0	117
						62	65.36 %
						34.64 %	
					1	288	736
						28.13 %	71.88 %
Comprobación	367	69.21 %	0.760	0.519	Comprobación	0	34
						16	68.00 %
						32.00 %	
					1	97	220
						30.60 %	69.40 %
Validación	169	72.78 %	0.755	0.510	Validación	0	19
						7	73.08 %
						26.92 %	
					1	39	104
						27.27 %	72.73 %

Fuente: Elaboración propia.

GRÁFICO 2. Importancia relativa de las variables de entrada en el árbol de decisión CHAID.



Fuente: Elaboración propia.

El Gráfico 2 muestra la importancia relativa de cada una de las variables predictoras significativas en el árbol de decisión CHAID. La importancia de estos predictores se determina calculando la reducción de la varianza de la variable objetivo (finalización o no de 4.º de la ESO por estudiante) que se puede atribuir a cada predictor mediante la realización de un análisis de sensibilidad (Saltelli et al., 2004). En esta investigación, la variable que acumula el 44 % de la importancia relativa es el número de años que ha repetido el estudiante (*years_repetition*), seguida de la variable que indica si el estudiante ha re-

petido o no (*student_repetition*) con una importancia del 30 %. A continuación, con un 16 %, se encuentra el tipo de programa en el que participó el estudiante (*kind_program*) y, finalmente, con un 10 % si el estudiante finalizó el programa especial de apoyo (*finish_program*).

3.2.2. Primera fase del análisis empírico: Análisis de los resultados del modelo árbol de decisión CHAID

El modelo ML basado en CHAID da lugar a un árbol en el que el primer grupo, denominado *nodo raíz*, representa el conjunto

total de la muestra, en este caso, de los 1739 estudiantes analizados. El algoritmo divide este conjunto en dos o más categorías denominadas *nodos padre o iniciales*. Debajo de los nodos padre se encuentran vinculados a ellos los nodos hijo. Las categorías que se encuentran en el último nivel del árbol de decisión se denominan *nodos terminales*. En lo que respecta a la jerarquía, los *nodos padre o iniciales* ejercen la mayor influencia sobre el *nodo raíz*, correspondiente a la variable dependiente que pretende explicar, mientras que ejercen una menor influencia sobre los *nodos terminales*.

Como se observa en el Gráfico 1, el número total de nodos es de 12, distribuidos de la siguiente forma en 3 niveles de profundidad: 1 nodo raíz (nivel 0), 4 nodos padre (distribuidos en los niveles 1 y 2) y 7 nodos terminales (distribuidos en los niveles 2 y 3). Además, en base al análisis de evaluación del modelo realizado en el apartado anterior se ha demostrado que este modelo tiene una buena capacidad de segmentación y predicción (ROC de 0.762).

3.2.2.1. Nivel 0 del Árbol de decisión CHAID

En el Nivel 0 del árbol de decisión CHAID, el Node 0 es la variable a predecir, que determina si el estudiante ha finalizado con éxito 4.º de la ESO (*finish_studies*). En este nivel se observa que, para el total de la muestra, el 85.34 % ha superado el curso, mientras que el 14.66 % no lo ha conseguido.

3.2.2.2. Nivel 1 del Árbol de decisión CHAID

En este nivel, la variable más relevante para definir los dos primeros nodos padre (Node 1 y Node 2) es el tipo de programa en el que el estudiante ha participado

(*kind_program*). Cada uno de estos nodos dará lugar a un subárbol que usarán diferentes variables predictoras para definir los siguientes niveles. En este nivel, la variable de mayor relevancia es el programa en el que participado el estudiante, lo llegue a finalizar o no. El 90.51 % de los estudiantes que ha participado en el programa C2 (acompañamiento del estudiante durante todo 4.º de la ESO) ha superado el curso 4º de la ESO, en comparación con el 68.61 % de aprobados para los que han participado en los programas C3 y C2C3. Por lo tanto, el programa con mayor éxito en sus objetivos es el C2.

3.2.2.3. Nivel 2 del Subárbol 1 (Izquierda) de decisión CHAID

La siguiente variable más relevante habiendo participado en el programa C2 es el número de veces que han repetido curso (*years_repetition*). De forma general, ya sabemos que en el Node 1 aprueban el 90.51 % de los estudiantes, pero este porcentaje puede variar mucho dependiendo de si el estudiante ha repetido algún curso o no previamente: los estudiantes que no han repetido curso aumentan su porcentaje de aprobado hasta el 96.46 % (Node 3), mientras que aquellos que han repetido una vez lo ven reducido hasta el 86.27 % (Node 4) y los que han repetido dos o más veces hasta el 77.59 % (Node 5). Los no repetidores obtienen mejores resultados en el curso 4º de la ESO en comparación con los que han repetido algún curso en alguna ocasión.

En este Nivel 2 encontramos los nodos Node 3, 4 y 5. Por un lado, los nodos Node 3 y 4, en los que los estudiantes no han repetido curso o lo han hecho solo una vez,

dan lugar a los nuevos subárboles del Nivel 3. En este Nivel 3 la variable más discriminante es la que determina si el estudiante llegó a finalizar el programa C2 en el que participó (`finish_program`). Por otro lado, Node 5 del Nivel 2 es un nodo terminal (no tiene más subniveles), lo que implica que la probabilidad de superar 4.º de la ESO para los estudiantes de este grupo es independiente de su desempeño en el programa C2, al contrario que para los nodos Node 3 y 4.

3.2.2.4. Nivel 3 del Subárbol 1 (Izquierda) de decisión CHAID

En el Nivel 3 se encuentran 4 nodos terminales: Node 8 y Node 9 corresponden a los estudiantes que habiendo participado en C2 no han repetido anteriormente y proceden del subárbol de Node 3. Los estudiantes del Node 9 han finalizado el programa C2 por lo que el 97.37 % han logrado superar el curso 4.º de la ESO, la cifra de mayor éxito de todos los segmentos del árbol de decisión CHAID. Mientras que en Node 8 no han llegado a finalizar el programa C2 y su porcentaje de éxito disminuye al 93.71 %.

Node 10 y Node 11 corresponden a los estudiantes en C2 que han repetido curso en una ocasión y proceden del subárbol de Node 4. Los estudiantes del Node 11 han finalizado el programa C2 por lo que el 88.06 % han logrado superar el curso 4.º de la ESO, mientras que los de Node 10 ven reducida esta cifra hasta el 80.39 % por no haber finalizado el programa C2. Por lo tanto, los estudiantes que finalizan el programa C2 obtienen mejores resultados en el curso 4.º de la ESO en comparación con los que no han llegado a finalizar el programa.

3.2.2.5. Nivel 2 del Subárbol 2 (Derecha) de decisión CHAID

En el nivel 2 del subárbol 2 (derecha, haber formado parte de C2C3 o C3) es si el estudiante ha repetido alguna vez curso (`student_repetition`). Sabemos que en el Node 2 aprueban tan solo el 68.61 % de los estudiantes, pero este porcentaje puede variar mucho dependiendo de si el estudiante es repetidor o no. Se obtiene que los estudiantes que no han repetido ningún curso aumentan su porcentaje de aprobado de 4.º de la ESO hasta el 78.67 % (Node 6), mientras que aquellos que han repetido curso ven reducido este porcentaje hasta el 58 % (Node 7). Por lo tanto, los no repetidores obtienen mejores resultados en el curso 4.º de la ESO en comparación con los que han repetido en alguna ocasión. En este subárbol 2 (derecha) ya no existen más niveles siendo esta variable la última relevante para generar más grupos de estudiantes.

3.3. Segunda fase del análisis empírico: análisis de robustez y calidad predictiva del modelo

En esta investigación se aplica una metodología de triangulación para comprobar la robustez de los resultados obtenidos en la primera fase. Para ello, se desarrolla un modelo predictivo ML basado en una red neuronal artificial perceptrón multicapa (ANN-MLP). Ambos modelos, utilizan la misma base de datos con las mismas variables de entrada (caracterización del estudiante, desempeño y tipo de programa realizado) y salida (desempeño del estudiante en 4.º de la ESO) (Tabla 1), pero están contruidos utilizando métodos de ML completamente diferentes, con el objetivo de verificar si obtienen resultados confirmatorios (Wolszczak-Derlacz y Parteka, 2011).

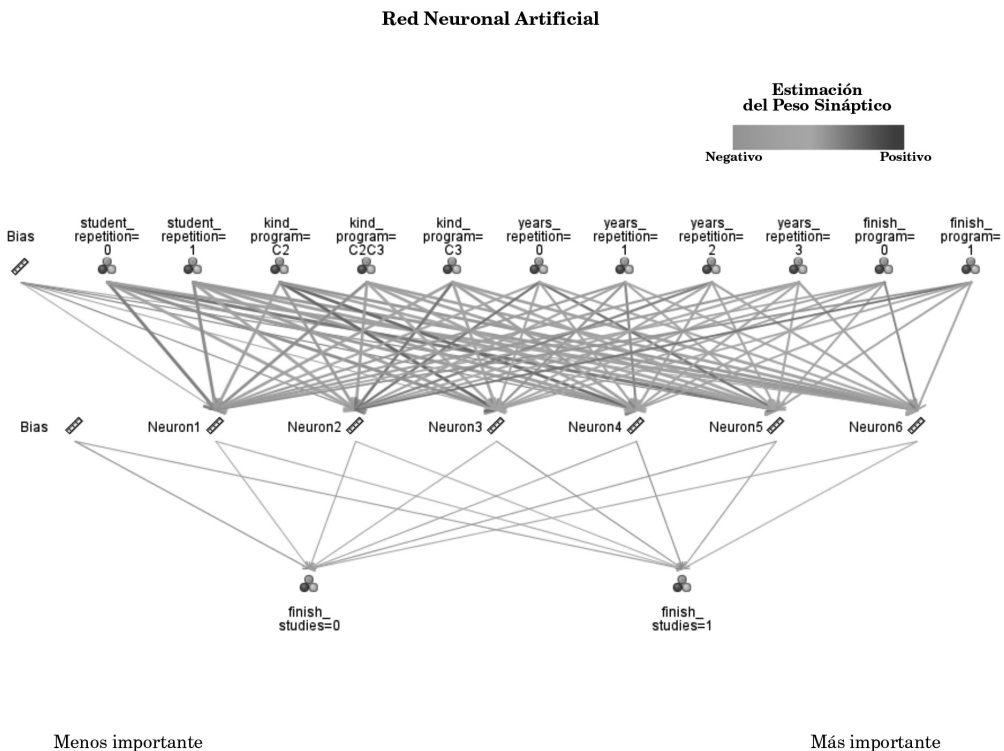
Se eligió ANN-MLP por la capacidad de las redes neuronales para manejar relaciones complejas entre variables, tanto lineales como no lineales, lo que las convierte en alternativa a modelos predictivos más tradicionales como regresiones logísticas donde la variable a predecir es dicotómica (Paliwal y Kumar, 2009). De esta forma, se asegura que no existen ningún tipo de relaciones que se hayan dejado de identificar.

A la hora de entrenar la ANN-MLP se debe tener en cuenta si la distribución de la variable dependiente es desbalanceada. Este sesgo en nuestra muestra podría impactar negativamente en el proceso de

entrenamiento de la ANN-MLP. Para evitarlo, se aplica un método de *oversampling* sobre el grupo menos representado para entrenar el modelo con el 70 % (2098 registros) de la muestra ya balanceada, garantizando que ambos grupos de estudiantes están igualmente representados a la hora de entrenar el modelo.

Por otro lado, la validación y comprobación del modelo ANN-MLP se realiza respectivamente con el 20 % (367 registros) y 10 % (169 registros) de la muestra, esta vez, sin balancear. Como en el primer modelo, se sigue también una metodología, entrenamiento, comprobación y validación (TTV).

GRÁFICO 3. Arquitectura de la ANN-MLP para predecir el éxito de los programas de apoyo a estudiantes para la superación del curso 4.º de la ESO.



Fuente: Elaboración propia.

El Gráfico 3 muestra la arquitectura utilizada para predecir el éxito de los programas de apoyo a estudiantes para la superación de 4.º de la ESO formada por tres capas: la primera capa de entrada consta de once unidades, correspondientes a los posibles valores de las cuatro variables de entrada relativos a las características de los estudiantes, su desempeño y el programa realizado (Tabla 1). La siguiente es la intermedia y se denomina capa oculta, consta de seis unidades o neuronas. Finalmente, la última consta de dos unidades, correspondientes a la predicción sobre el desempeño del estudiante en 4.º de la ESO (Tabla 1). El tipo de función de activación de la capa oculta y la capa de salida corresponden con una tangente hiperbólica y *softmax* respectivamente.

En la Tabla 3 se describen los indicadores de precisión y la matriz de confusión de la ANN-MLP que demuestran que la calidad de este modelo es buena: clasificación correcta del 71.19 % de los estudiantes y un AUC de 0.763 (Hosmer Jr. et al., 2013). Además, también se verifica que no existe sobre-entrenamiento ya que los indicadores obtenidos tanto para el total de la muestra como para las submuestras de entrenamiento, comprobación y validación son similares.

4. Discusión

Los indicadores que hemos utilizado para evaluar la precisión de la ANN-MLP como predictor del éxito de los estudiantes son los mismos que los utilizados para el CHAID, facilitando así la comparación entre resultados, calidad en la predicción y robustez de ambos modelos.

En la Tabla 4 se realiza la comparativa entre los indicadores de la precisión en la clasificación y matriz de confusión de ambos modelos (Tabla 3 vs. Tabla 2). Se observa que las predicciones y precisión de la ANN-MLP son consistentes con las realizadas por el árbol de decisión CHAID. La ANN-MLP predice correctamente el 71.19 % (error del 8.81 %) de los estudiantes que superarán el curso de 4.º de la ESO. Esta cifra representa un 0.46 % (8 estudiantes) de mayor capacidad predictiva frente al 70.73 % de aciertos obtenidos por el árbol de decisión CHAID. Para el resto de los indicadores, como la sensibilidad, especificidad, falsos positivos y falsos negativos, las diferencias también son estadísticamente no significativas, oscilando entre -1.57 % (diferencia en la clasificación de 4 estudiantes) y 0.81 % (diferencia en la clasificación de 12 estudiantes). Lo mismo sucede con las diferencias existentes entre las áreas debajo de la curva (AUC) ROC y coeficiente de GINI. El AUC para la ANN-MLP y el árbol de decisión CHAID, 0.763 y 0.762 respectivamente y el coeficiente de GINI 0.525 y 0.523 respectivamente. En el Gráfico 4 se muestra la importancia relativa de las variables o predictores de entrada de la ANN-MLP siguiendo el método de la varianza y en total suman 1.

Si bien es cierto que tanto el árbol de decisión CHAID como la ANN-MLP utilizan la misma base de datos y variables en el modelo y consiguen alcanzar resultados, capacidad de predicción y robustez muy similares, la forma de hacerlo es muy diferente, tanto en los métodos utilizados como en el manejo e importancia de las variables que contribuyen en el modelo.

Precisión del Modelo

Precisión del Modelo			Matriz de Confusión									
			Predicción									
Muestra	Tamaño Muestra	Porcentaje Correcto	AUC	GINI	Muestra	Observado	Tamaño muestra			Porcentaje		
Total Muestra	1739	71.19 %	0.763	0.525	Total Muestra	0	166	89	65.10 %	34.90 %		
						1	412	1072	27.76 %	72.24 %		
Submuestra												
Entrenamiento	2098*	69.78 %	0.767	0.535	Entrenamiento	0	720	354	67.04 %	32.96 %		
						1	280	744	27.34 %	72.66 %		
Comprobación	367	70.03 %	0.755	0.510	Comprobación	0	31	19	62.00 %	38.00 %		
						1	91	226	28.71 %	71.29 %		
Validación	169	69.23 %	0.744	0.488	Validación	0	15	11	57.69 %	42.31 %		
						1	41	102	28.67 %	71.33 %		

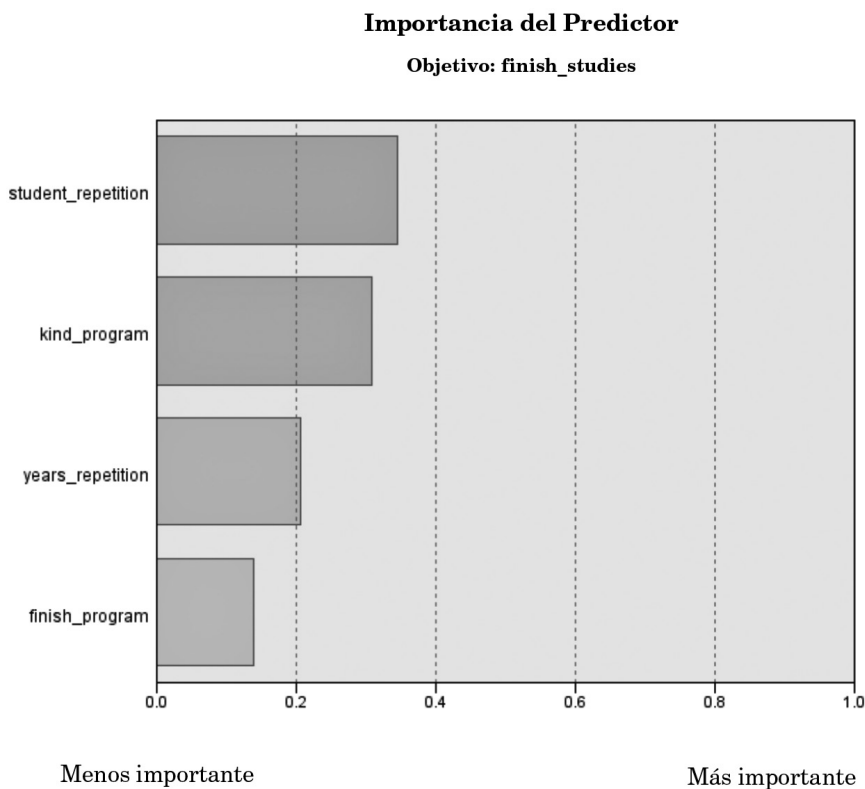
Fuente: Elaboración propia.

TABLA 4. Comparación indicadores: Precisión de la clasificación y matriz de confusión de la ANN-MLP en comparación con el árbol de decisión CHAID.

Indicadores	ANN-MLP comparado con Arbol CHAID	
	Variación %	Variación clasificados
Porcentaje Correcto	0.46 %	8
Sensibilidad	0.81 %	12
Especificidad	-1.57 %	-4
Falsos positivos	1.57 %	4
Falsos negativos	-0.81 %	-12
AUC	0.001	-
GINI	0.002	-

Fuente: Elaboración propia.

GRÁFICO 4. Importancia relativa de las variables de entrada de la red neuronal artificial perceptrón multicapa (ANN-MLP).



Fuente: Elaboración propia.

En el caso de la ANN-MLP la variable que acumula mayor importancia relativa acumula hasta el 35 % y es la que indica si el estudiante ha repetido algún curso (*student_repetition*), seguida del tipo de programa en el que participó el estudiante con una importancia del 31 % (*kind_program*). A continuación, con una importancia del 21 % se encuentra la variable que recoge el número de cursos que ha repetido el estudiante (*years_repetition*) y, finalmente, con una importancia del 13 %, la variable que expresa si el estudiante finalizó el programa especial de apoyo (*finish_program*).

La aplicación de la metodología de triangulación concluye que los modelos de árbol de decisión CHAID y la ANN-MLP obtienen resultados similares en cuanto a precisión de clasificación y predicción del éxito de los programas de apoyo a estudiantes para la superación de 4.º de la ESO y, por lo tanto, son modelos robustos que confirman sus resultados. De los tres programas analizados el que mayor porcentaje de éxito ha conseguido es el C2 que acompaña al estudiante durante todo el curso de 4.º de la ESO, alcanzando un 90.51 % de aprobados del curso, comparado con el 68.61 % de las otras dos opciones C2 (clases extraordinarias en el mes de Julio) y C2C3 (combinación de las opciones anteriores).

En términos de evaluación educativa los resultados son impactantes: se puede evaluar el retorno económico y social de una intervención en función de las distintas características de los estudiantes, como el aprovechamiento de los cursos, el tipo de curso o, por supuesto, si son repetidores o no.

5. Conclusiones

A través de nuestra investigación hemos contrastado la necesidad de evaluar económicamente la LOMLOE. La ley partía con la misma dotación presupuestaria que la LOMCE para intervenciones como la analizada (45 millones de euros). La llegada de fondos europeos abre nuevas posibilidades no contempladas en su diseño inicial, por lo que se abre el reto de emplear esa inversión adicional de forma eficiente.

Nuestro análisis demuestra que se puede predecir la probabilidad que tienen los estudiantes en riesgo de superar 4.º de la ESO en función del programa de apoyo en el que participan. De esta forma se puede contribuir al diseño de programas más eficientes y con mayores tasas de éxito.

Para completar los modelos ya utilizados por Administraciones públicas, empleamos modelos de ML robustos como árboles de decisión CHAID y redes neuronales artificiales para analizar las características de los estudiantes, la intervención en la que han formado parte y los resultados finales. Esto permite maximizar el retorno social de cada programa y apoyar a aquellos estudiantes con menos posibilidades de éxito.

Por ejemplo, nuestros resultados muestran que, en el *Programa para la Mejora del Éxito Educativo* en Castilla y León, independientemente del programa en el que el estudiante ha participado, aquellos que nunca han repetido tienen mayor probabilidad de aprobar 4.º de la ESO que los que lo han hecho. La probabilidad de superar el curso aumenta cuando los estudiantes, además de participar en los programas especia-

les, lo finalizan. Por este motivo, el grupo de estudiantes que alcanza mejores resultados es el de no repetidores que participan y finalizan el programa C2 con un 97.37 % de aprobados de 4.º de la ESO. Mientras que el que peores resultados obtiene es el de repetidores que han participado en los programas C2C3 y C3 independientemente de que lo hayan finalizado o no con un 58 % de aprobados en 4.º de la ESO. Nuestro análisis también muestra que realizar los apoyos educativos en julio no constituyen un gran valor añadido por cuanto que el programa C1 con acompañamiento durante el curso obtiene mejores resultados que el C3 que aúna este refuerzo durante el año escolar y uno añadido en el mes de julio. En la siguiente fase de investigación comprobaremos si estos diferentes resultados se deben a los propios alumnos o es por las características del programa.

Como corolario, también demostramos que los modelos de ML como CHAID y redes neuronales artificiales son buenos candidatos a la hora de construir modelos que den respuesta a las preguntas educativas planteadas, como la importancia de tutelar al estudiante para que finalice dicho programa y no lo abandone, especialmente en el caso de los jóvenes que han repetido para los que el programa supone una herramienta muy eficiente para evitar el fracaso escolar y una forma eficiente de inversión del dinero público.

Nota

¹ <https://educationendowmentfoundation.org.uk/projects-and-evaluation/projects/the-impact-of-covid-19-on-school-starters/>

Referencias bibliográficas

- Aad, G., Abbott, B., Abdallah, J., Abdelalim, A. A., Abdesselam, A., Abidinov, O., Abi, B., Abolins, M., Abramowicz, H., Abreu, H., Acerbi, E., Acharya, B. S., Adams, D. L., Addy, T. N., Adelman, J., Aderholz, M., Adomeit, S., Adragna, P., Adye, T., ... Zwailinski, L. (2012). Performance of missing transverse momentum reconstruction in proton-proton collisions at $\sqrt{s} = 7$ TeV with atlas [Rendimiento de la reconstrucción del momento transversal perdido en colisiones protón-protón en $\sqrt{s} = 7$ TeV con atlas]. *European Physical Journal C*, 72 (1), 1-35. <https://doi.org/10.1140/epjc/s10052-011-1844-6>
- Akin, M., Eydurán, E. y Reed, B. M. (2017). Use of RSM and CHAID data mining algorithm for predicting mineral nutrition of hazelnut [Uso del algoritmo de minería de datos RSM y CHAID para predecir la nutrición mineral de la avellana]. *Plant Cell, Tissue and Organ Culture (PCTOC)*, 128 (2), 303-316.
- Ballestar, M. T., Doncel, L. M., Sainz, J. y Ortigosa-Blanch, A. (2019). A novel machine learning approach for evaluation of public policies: An application in relation to the performance of university researchers [Un novedoso enfoque de aprendizaje automático para la evaluación de políticas públicas: una aplicación en relación con el rendimiento de los investigadores universitarios]. *Technological Forecasting and Social Change*, 149. <https://doi.org/10.1016/j.techfore.2019.119756>
- Ballestar, M. T., García-Lázaro, A. y Sainz, J. (2020). Todos los caminos llevan a la educación: un primer análisis de la robotización, la educación y el empleo. *Papeles de Economía Española*, 166, 33-49.
- Ballestar, M. T., García-Lázaro, A., Sainz, J. y Sanz, I. (2022). Why is your company not robotic? The technology and human capital needed by firms to become robotic [¿Por qué su empresa no está robotizada? La tecnología y el capital humano que necesitan las empresas para robotizarse]. *Journal of Business Research*, 142, 328-343. <https://doi.org/10.1016/j.jbusres.2021.12.061>
- Ballestar, M. T., Grau-Carles, P. y Sainz, J. (2018). Customer segmentation in e-commerce: Applications to the cashback business model [Segmentación de clientes en el comercio electrónico: Aplicaciones al modelo de negocio

- del cashback]. *Journal of Business Research*, 88, 407-414. <https://doi.org/10.1016/j.jbusres.2017.11.047>
- Brunello, G. y de Paola, M. (2014). The costs of early school leaving in Europe [Los costes del abandono escolar en Europa]. *IZA Journal of Labor Policy*, 3 (1), 1-31. <https://doi.org/10.1186/2193-9004-3-22>
- Burgess, S. (16 de junio de 2020). *How we should deal with the lockdown learning loss in England's schools [Cómo debemos afrontar la pérdida de aprendizaje en las escuelas de Inglaterra]*. VOX Cepr Policy Portal. <https://bit.ly/3Eom12M>
- Chassignol, M., Khoroshavin, A., Klimova, A. y Bilyatdinova, A. (2018). Artificial intelligence trends in education: A narrative overview [Tendencias de la inteligencia artificial en la educación: una visión general]. *Procedia Computer Science*, 136, 16-24. <https://doi.org/10.1016/j.procs.2018.08.233>
- Chatterjee, S. y Bhattacharjee, K. K. (2020). Adoption of artificial intelligence in higher education: A quantitative analysis using structural equation modelling [Adopción de la inteligencia artificial en la enseñanza superior: un análisis cuantitativo mediante un modelo de ecuaciones estructurales]. *Education and Information Technologies*, 25 (5), 3443-3463. <https://doi.org/10.1007/s10639-020-10159-7>
- Crescenzi, R., Giua, M. y Sonzogno, G. V. (2021). Mind the Covid-19 crisis: An evidence-based implementation of Next Generation EU. *Journal of Policy Modeling*, 43 (2), 278-297. <https://doi.org/10.1016/J.JPOLMOD.2021.03.002>
- Cruz-Jesús, F., Castelli, M., Oliveira, T., Mendes, R., Nunes, C., Sa-Velho, M. y Rosa-Louro, A. (2020). Using artificial intelligence methods to assess academic achievement in public high schools of a European Union country. *Heliyon*, 6 (6), E04081. <https://doi.org/10.1016/j.heliyon.2020.e04081>
- De la Fuente, A., & Jimeno, J. F. (2009). The private and fiscal returns to schooling in the European Union [Los rendimientos privados y fiscales de la escolarización en la Unión Europea]. *Journal of the European Economic Association*, 7 (6), 1319-1360. <https://doi.org/10.1162/JEEA.2009.7.6.1319>
- Doncel, L. M., Durá, P., Grau, P. y Sainz, J. (2014). Fiscal sustainability and immigration in the Madrid region [Sostenibilidad fiscal e inmigración en la Comunidad de Madrid]. *International Migration*, 52 (6), 180-196. <https://doi.org/10.1111/imig.12004>
- Dželihodžić, A. y Jonko, D. (2016). Comparison of ensemble classification techniques and single classifiers performance for customer credit assessment [Comparación de las técnicas de clasificación por conjuntos y el rendimiento de los clasificadores individuales para la evaluación del crédito de los clientes]. *Modeling of Artificial Intelligence*, 3, 140-150.
- European Commission (2010). *EUROPE 2020: A strategy for smart, sustainable and inclusive growth [EUROPA 2020: Una estrategia para un crecimiento inteligente, sostenible e integrador]*. <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2010:2020:FIN:EN:PDF>
- Golden, G. (2020). Education policy evaluation: Surveying the OECD landscape [Evaluación de la política educativa: un estudio del panorama de la OCDE]. *OECD Education Working Papers*, 236. OECD Publishing. https://www.oecd-ilibrary.org/education/education-policy-evaluation_9f127490-en
- Goos, M., Manning, A. y Salomons, A. (2009). Job polarization in Europe [Polarización del empleo en Europa]. *American Economic Review*, 99 (2), 58-63. <https://doi.org/10.1257/aer.99.2.58>
- Gouëdard, P., Pont, B. y Viennet, R. (2020). Education responses to COVID-19: Implementing a way forward. *OECD Education Working Papers*, 224. OECD Publishing. <https://doi.org/https://doi.org/10.1787/8e95f977-en>
- Gregory, T., Salomons, A. y Zierahn, U. (2019). Racing With or Against the Machine? Evidence from Europe. *IZA Institute of Labor Economics*, 2019 (12063).
- Hosmer Jr., D. W., Lemeshow, S. y Sturdivant, R. X. (2013). *Applied logistic regression [Regresión logística aplicada]*. John Wiley & Sons.
- Kass, G. V. (1980). An exploratory technique for investigating large quantities of categorical data [Una técnica exploratoria para investigar grandes cantidades de datos categóricos]. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 29 (2), 119-127. <https://doi.org/10.2307/2986296>
- Khosravi, P., Kazemi, E., Zhan, Q., Malmsten, J. E., Toschi, M., Zisimopoulos, P., Sigaras, A., La-

- very, S., Cooper, L. A. D. y Hickman, C. (2019). Deep learning enables robust assessment and selection of human blastocysts after in vitro fertilization [El aprendizaje profundo permite evaluar y seleccionar de forma robusta los blastocistos humanos tras la fecundación in vitro]. *NPJ Digital Medicine*, 2 (1), 1-9.
- Kraft, M. (20 de agosto de 2015). What If Every Struggling Student Had a Tutor? [¿Y si todos los estudiantes con dificultades tuvieran un tutor?] *Education Week*. <https://www.edweek.org/education/opinion-what-if-every-struggling-student-had-a-tutor/2015/08>
- Ley Orgánica 10/2002, de 23 de diciembre, de Calidad de la Educación (LOCE). *Boletín Oficial del Estado*, 307, de 24 de diciembre de 2002, páginas 45188 a 45220. <https://www.boe.es/eli/es/lo/2002/12/23/10>
- Ley Orgánica 2/2006, de 3 de mayo de Educación (LOE). *Boletín Oficial del Estado*, 106. <https://www.boe.es/eli/es/lo/2006/05/03/2/con>
- Ley Orgánica 8/2013, de 9 de diciembre, para la Mejora de la Calidad Educativa (LOMCE). *Boletín Oficial del Estado*, 295, páginas 97858 a 97921. <https://www.boe.es/eli/es/lo/2013/12/09/8>
- Ley Orgánica 3/2020, de 29 de diciembre, por la que se modifica la Ley Orgánica 2/2006, de 3 de mayo, de Educación (LOMLOE). *Boletín Oficial del Estado*, 340, de 30 de diciembre de 2020, páginas 122868 a 122953. <https://www.boe.es/eli/es/lo/2020/12/29/3>
- Li, Z., & Eastman, J. R. (2006). Commitment and typicality measurements for fuzzy ARTMAP neural network [Medidas de compromiso y tipicidad para la red neuronal difusa ARTMAP]. *Geoinformatics 2006: Geospatial Information Science*, 6420. <https://doi.org/10.1117/12.712998>
- Maimon, O. y Rokach, L. (2005). *Data mining and knowledge discovery handbook [Manual de minería de datos y descubrimiento de conocimientos]*. Springer.
- Nickow, A., Oreopoulos, P. y Quan, V. (2020). The impressive effects of tutoring on prek-12 learning: A systematic review and meta-analysis of the experimental evidence [Los impresionantes efectos de la tutoría en el aprendizaje de preescolar a 12 años: una revisión sistemática y un meta-análisis de las pruebas experimentales]. *National Bureau of Economic Research Working Papers Series*, 27476. <https://doi.org/10.3386/w27476>
- OECD. (2018). *Education policy outlook 2018: Putting student learning at the centre [Perspectivas de la política educativa 2018: Poner el aprendizaje de los estudiantes en el centro]*. OECD.
- Paliwal, M. y Kumar, U. A. (2009). Neural networks and statistical techniques: A review of applications [Redes neuronales y técnicas estadísticas: una revisión de las aplicaciones]. *Expert Systems with Applications*, 36 (1), 2-17.
- Pont, B. y Montt, G. (2014). *Education policy outlook. Spain [Perspectivas de la política educativa. España]*. OECD.
- Porte, C. de la y Jensen, M. D. (2021). The next generation EU: An analysis of the dimensions of conflict behind the deal [La UE de nueva generación: un análisis de las dimensiones del conflicto tras el acuerdo]. *Social Policy & Administration*, 55 (2), 388-402. <https://doi.org/10.1111/SPOL.12709>
- Ramaswami, M. y Bhaskaran, R. (2010). A CHAID based performance prediction model in educational data mining. *International Journal of Computer Science*, 7 (1), 10-18.
- Saltelli, A., Tarantola, S., Campolongo, F. y Ratto, M. (2004). *Sensitivity analysis in practice: a guide to assessing scientific models (Vol. 1)*. Wiley Online Library.
- Schleicher, A. (2020). The impact of COVID-19 on education: Insights from education at a glance 2020 [El impacto de COVID-19 en la educación: visión de la educación en un vistazo 2020]. *OECD Journal: Economic Studies*. <https://www.oecd.org/education/the-impact-of-covid-19-on-education-insights-education-at-a-glance-2020.pdf>
- Slavin, R. E. (2016). Evidence-based education policies: Transforming educational practice and research [Políticas educativas basadas en la evidencia: transformar la práctica y la investigación educativas]. *Educational Researcher*, 31 (7), 15-21. <https://doi.org/10.3102/0013189X031007015>
- Wolszczak-Derlacz, J. y Parteka, A. (2011). Efficiency of European public higher education institutions: A two-stage multicountry approach [Eficiencia de las instituciones públicas europeas de enseñanza superior: un enfoque multinacio-

- nal en dos fases]. *Scientometrics*, 89 (3), 887-917. <https://doi.org/10.1007/s11192-011-0484-9>
- Yin, L., Ge, Y., Xiao, K., Wang, X. y Quan, X. (2013). Feature selection for high-dimensional imbalanced data [Selección de características para datos desequilibrados de alta dimensión]. *Neurocomputing*, 105, 3-11.
- Yusuf, A. (2007). *Standard-setting at UNESCO: Conventions, recommendations, declarations and charters adopted by UNESCO, 1948-2006 (Vol. 2) [La elaboración de normas en la UNESCO: Convenciones, recomendaciones, declaraciones y cartas adoptadas por la UNESCO, 1948-2006 (Vol. 2)]*. Martinus Nijhoff Publishers.

Biografía de los autores

María Teresa Ballestar es Graduada en Estadística por la Universidad de Zaragoza, Licenciada en Investigación y Técnicas de Mercado y Master en Sociedad de la Información y el Conocimiento en la Universitat Oberta de Catalunya. Doctorada en metodologías y tecnologías Big Data aplicadas a la Economía en la Universidad Rey Juan Carlos. Durante los últimos años ha ocupado puestos de dirección y liderado proyectos de transformación digital, innovación, data analytics y data science. Además, es Profesora Asociada de la Universidad Rey Juan Carlos y ha publicado más de una decena de artículos científicos y de divulgación y colaborado con distintos medios de comunicación.



<https://orcid.org/0000-0001-8526-7561>

Jorge Sainz Orcid es Licenciado en Ciencias Económicas por la Universidad Complutense de Madrid y en Derecho por la UNED. Doctor en Economía (URJC) y

MBA (especialización en finanzas y políticas públicas) Simon School, University of Rochester. Es Catedrático de Economía Aplicada en la Universidad Rey Juan Carlos de Madrid y Visiting Fellow en el Institute for Policy Research, University of Bath.

Ha sido Asesor del Gabinete de la consejera de Educación y Subdirector General de Investigación de la Consejería de Educación de la Comunidad de Madrid. En el Ministerio de Educación, Cultura y Deporte ha sido director general de Política Universitaria y Secretario General de Universidades.



<https://orcid.org/0000-0001-8491-3154>

Ismael Sanz es Profesor Titular del Departamento de Economía Aplicada I de la Universidad Rey Juan Carlos. Doctor en Economía Aplicada en la Facultad de Ciencias Económicas y Empresariales de la Universidad Complutense de Madrid (UCM). Ha sido Director General de Innovación, Becas y Ayudas de la Consejería de Educación de la Comunidad de Madrid, Director del Instituto Nacional de Evaluación Educativa (INEE) del Ministerio de Educación, Cultura y Deporte (2012-2015) y Chair del Strategic Development Group de PISA de la OCDE (noviembre 2014-julio de 2015). Actualmente es Vicerrector de Calidad de la URJC. Autor de artículos en revistas de investigación de alto impacto.



<https://orcid.org/0000-0003-1286-4124>

Sumario *

Table of Contents **

La LOMLOE ante los desafíos de la educación española en el siglo XXI *The LOMLOE amidst the challenges of the Spanish education system in the 21st century*

Editor invitado: Francisco López Rupérez
Guest editor: Francisco López Rupérez

Francisco López Rupérez

Presentación: La LOMLOE ante los desafíos de la educación española en el siglo XXI
Introduction: The LOMLOE amidst the challenges of the Spanish education system in the 21st century

3

Estudios y ensayos *Studies and essays*

Francisco Esteban Bara y Fernando Gil Cantero

Las finalidades de la educación y la LOMLOE: cuestiones controvertidas en la acción educativa
The purposes of education and Spain's LOMLOE educational legislation: Controversial questions in educational action

13

José Luis Gaviria y David Reyero

La transmisión de los contenidos culturales y su evaluación entre los fines del sistema educativo, según la LOMLOE
The transmission of cultural content and its evaluation among the ends of the education system: An analysis of the LOMLOE

31

Francisco López Rupérez

El enfoque del currículo por competencias.
Un análisis de la LOMLOE
The transmission of cultural content and its evaluation among the ends of the education system: An analysis of the LOMLOE

55

José Luis Martínez López-Muñiz

La calidad regulatoria de los marcos normativos: una aproximación crítica
The regulatory quality of legal frameworks: A critical approach

69

Charles L. Glenn

Pluralismo educativo y niños vulnerables
Educational pluralism and vulnerable children

85

José Adolfo de Azcárraga

La nueva legislación educativa: por qué no mejorará la educación pública en España
The new Spanish educational legislation: Why public education will not improve

111

* Todos los artículos están también publicados en inglés en la página web de la revista: <https://revistadepedagogia.org/en>.

** All the articles are also published in English on the web page of the journal: <https://revistadepedagogia.org/en>.

Investigaciones y propuestas

Testing and proposals

María Teresa Ballestar, Jorge Sainz e Ismael Sanz

Evaluación económica de intervenciones educativas en la LOMLOE: propuestas de mejora con inteligencia artificial
An economic evaluation of educational interventions in the LOMLOE: Proposals for improvement with Artificial Intelligence

133

Francisco López Rupérez

La calidad de la gobernanza del sistema educativo. El caso de la LOMLOE
The quality of governance of the education system. The case of the LOMLOE

155

Inmaculada Egido Gálvez

La reforma del currículo para responder a los retos del futuro. España en perspectiva internacional
Curriculum reform to respond to the challenges of the future. Spain in international perspective

175

Samuel Gento Palacios, Raúl González-Fernández y Ernesto López-Gómez

Dirección de centros educativos y ampliación de autonomía con rendición de cuentas. El rol mediador del liderazgo pedagógico
Heads of educational institutions and expansion of autonomy with accountability. The mediating role of pedagogical leadership

193

Reseñas bibliográficas

Escámez-Sánchez, J. y Peris-Cancio J.-A. (2021).

La universidad del siglo xxi y la sostenibilidad social (Ramón Mínguez-Vallejos). **Grupo SI(e)TE Educación (2021).** La calidad en la educación (José Antonio Jordán). **Santos-Rego, M. A., Lorenzo, M. y Mella, I. (2020).** *El aprendizaje-servicio y la educación universitaria. Hacer personas competentes* (Alexandre Sotelino Losada).

211

Informaciones

Congreso anual de Filosofía de la Educación de la Sociedad de Filosofía de la Educación de Gran Bretaña; IX Congreso Internacional «La Palabra en la Educación. El papel del arte, la literatura y la belleza en el fomento de la resiliencia y el crecimiento postraumático»; XX Congreso Internacional de Investigación Educativa «Educación inclusiva y equitativa de calidad en tiempos de crisis»; ECER Conference Education in a Changing World: The impact of global realities on the prospects and experiences of educational research.

223



ISSN: 0034-9461 (Impreso), 2174-0909 (Online)

<https://revistadepedagogia.org/>

Depósito legal: M. 6.020 - 1958

INDUSTRIA GRÁFICA ANZOS, S.L. Fuenlabrada - Madrid

An economic evaluation of educational interventions in the LOMLOE*: Proposals for improvement with Artificial Intelligence

Evaluación económica de intervenciones educativas en la LOMLOE: Propuestas de mejora con Inteligencia Artificial

María Teresa BALLESTAR, PhD. Associate Professor. Universidad Rey Juan Carlos (teresa.ballestar@urjc.es).

Jorge SAINZ, PhD. Professor. Universidad Rey Juan Carlos (jorge.sainz@urjc.es).

Ismael SANZ, PhD. Associate Professor. Universidad Rey Juan Carlos (ismael.sanz@urjc.es).

Abstract:

This research aims to demonstrate the need for an economic evaluation of the Organic Law that modifies the Organic Law of Education (LOMLOE), especially after the investment of EU Next Generation funds that open new opportunities that were lacking in the initial drafting of the law. The challenge for Public Administrations is to use this additional investment efficiently.

Our analysis shows that artificial intelligence models can predict whether educational support programmes will help increase the likelihood that students who lag behind will pass

the 4th grade of ESO (Compulsory Secondary Education). In this way, we can calculate the social return of one of these programmes and contribute to their *ex-ante* design to achieve higher success rates for students.

To complement the models already used by Public Administrations, we use robust Machine Learning (ML) models such as CHAID decision trees and artificial neural networks to analyse the characteristics of the groups of students and the intervention they have been part of. The conclusions allow us to improve educational reinforcement programmes in the coming years to

* Organic Law 3/2020, of 29 December, which amends Organic Law 2/2006, of 3 May, on Education (LOMLOE).

Revision accepted: 2021-12-13.

This is the English version of an article originally printed in Spanish in issue 281 of the **revista española de pedagogía**. For this reason, the abbreviation EV has been added to the page numbers. Please, cite this article as follows: Ballestar, M. T., Sainz, J., & Sanz, I. (2022). Evaluación económica de intervenciones educativas en la LOMLOE: propuestas de mejora con inteligencia artificial | *An economic evaluation of educational interventions in the LOMLOE: Proposals for improvement with artificial intelligence*. *Revista Española de Pedagogía*, 80 (281), 133-154. <https://doi.org/10.22550/REP80-1-2022-09>

<https://revistadepedagogia.org/>

ISSN: 0034-9461 (Impreso), 2174-0909 (Online)

support students with lower chances of academic success.

Keywords: Public policy analysis, Machine Learning, educational efficiency, LOMLOE.

Resumen:

El objetivo de esta investigación es demostrar la necesidad de evaluar económicamente la LOMLOE (Ley Orgánica 3/2020, de 29 de diciembre de 2020, por la que se modifica la Ley Orgánica 2/2006, de 3 de mayo, de Educación), especialmente tras la inversión de los fondos EU Next Generation que abren nuevas oportunidades de las que carecía la ley en su redacción inicial. Las Administraciones públicas tienen el reto de emplear esa inversión adicional de forma eficiente.

Nuestro análisis demuestra que los modelos de inteligencia artificial pueden predecir si

los programas de apoyo educativo ayudarán a incrementar la probabilidad de que estudiantes rezagados superen 4.º de la ESO. De esta forma, se puede calcular el retorno social de los programas de apoyo educativo y contribuir a su diseño *ex-ante* para lograr que los alumnos tengan mayores tasas de éxito.

Para completar los modelos ya utilizados por Administraciones públicas, empleamos modelos de Machine Learning (ML) robustos como árboles de decisión CHAID y redes neuronales artificiales para analizar las características de los grupos de estudiantes y la intervención en la que han formado parte. Las conclusiones permiten mejorar los programas de refuerzo educativo de los próximos años para apoyar a los alumnos con menos posibilidades de éxito académico.

Descriptores: análisis de políticas públicas, *Machine Learning*, eficiencia educativa, LOMLOE.

1. Introduction

One of the principles of educational management by international organisations such as the OECD (Organisation for Economic Co-operation and Development) is the evaluation of educational policies, both global and of specific interventions. (Golden, 2020). The result of the analysis leads to greater effectiveness and impact, both at the social and individual levels, of educational actions. This in turn has a positive impact on students and a positive dynamic for the system in general. (OECD, 2018). This informative principle appears in almost all European law, and,

in the case of Spain, is included in Organic Law 2/2006 of 3 May (LOE), which states in Article 2 bis that “The functioning of the Spanish Education System is governed by the principles of [...] efficiency in the allocation of public resources, transparency and accountability”.

Our aim is to demonstrate that the economic evaluation of education policies serves to optimise and prioritise the legislative objectives set by the LOMLOE. This fact would demonstrate that the law faces a gap in its construction that affects its management evaluation, an evaluation

that is a requirement for the justification of *EU Next Generation* funds for education, which in 2021 alone represent 1,852.5 million euros in the general state budget.

It is this European investment that will make it possible to implement and evaluate the law, since in its Analysis and Regulatory Impact Report (known as MAIN), which includes the technical aspects of the regulation, it is stated that "...the project does not involve an increase or decrease in public spending". On the other hand, the Commission requires an assessment of whether the population effectively acquires the skills to compete at a global level. (Crescenzi et al., 2021; Porte & Jensen, 2021).

The European Commission and UNESCO recognise the importance of this type of action, both in the choice of reforms and in their subsequent analysis, and the need for data on educational interventions to be available for said analysis, despite the reluctance of some administrations. European Commission (EC, 2010; Yusuf, 2007). Although the preferred approach for this analysis is the experimental method, in the reality of education there are problems in carrying it out for various reasons: such as the economic cost, the difficulty of establishing randomised control groups, problems of management, anxiety, etc. (Golden, 2020; Slavin, 2016).

To solve this proposition, there is an arsenal of new techniques that are being used and which are in full development and can serve as a proposition, including artificial intelligence (AI) (Ballestar et al.,

2019; Chassignol et al., 2018; Chatterjee & Bhattacharjee, 2020) because of the difficulty of identifying who are the main beneficiaries and what are the long-term effects. Still, new policies including financial incentives have been adopted to increase the research output at all possible levels. Little literature has been devoted to the response to those incentives. To bridge this gap, we carry out our analysis with data of a six years program developed in Madrid (Spain). Following the triangulation strategy of Ballestar et al., (2020) we evaluate the use of different AI strategies for the evaluation of the return on educational interventions. To do so, we will delve into the intervention carried out by the Regional Ministry of Education of the Junta de Comunidades de Castilla y León called *Programa para la Mejora del Éxito Educativo (Programme for the Improvement of Educational Achievement)*, which we will analyse in its 2020 edition, after the closure of schools due to the pandemic. The Junta made a call for financial aid so that schools could offer support classrooms during July to students in the sixth year of primary school, the fourth year of ESO and the second year of baccalaureate in the areas of Spanish language and literature, mathematics, or English with educational difficulties. The programme has contributed over the years to the promotion of a significant number of students in the community. All grades benefited from the programme, yielding an improvement in promotion of 5% in the 2019-20 academic year for students who participated in the programme, compared to those who were able to participate but did not.

2. The intervention

To design any educational intervention Slavin (2016) y Golden (2020) recognise as fundamental that it should be based on scientific evaluations. The support classes that make up the *Programa para la Mejora del Éxito Educativo en Castilla y León* are probably one of the most scientifically grounded in recent years. Research conducted by the *Education Endowment Foundation*¹ shows that small group tutoring is among the measures for which there is empirical evidence of greatest effectiveness internationally. As such, they have proven to be a good complement to an education system that is designed to move large numbers of students from grade to grade, but which does not work for all students.

Kraft (2015) and Burgess (2020) discuss the factors contributing to its success: delivery by selected and trained staff in coordination with their regular teachers and in small groups. The impact is significant: Nickow et al., (2020) show, from the study of 96 randomised experiment items, that the effect of small-group tutoring is large and significant (37% of the standard deviation).

This intervention is one of the few that are included in the Spanish legislation. LOMLOE, Art. 4.4, states as one objective to facilitate "...access to the support that students require" and the increase of these measures throughout the entire educational process in an individualised manner (Artart. 20 bis) "...to avoid school repetition, particularly in socially disadvantaged environments".

In fact, the regulation leaves a very clear wording on its objectives to repetition, for example, the new Artart. 28 reads: "Remaining in the same year [of secondary school] shall be considered an exceptional measure and shall be taken after having exhausted the ordinary reinforcement and support measures to solve learning difficulties".

These support measures have been present in educational legislation throughout this century. Thus, the unborn Organic Law 10/2002, of 23 December, on the Law on the Quality of Education (LOCE) included in its art. 2 the right of students to "...receive the necessary help and support to compensate for personal, family, economic, social, and cultural deficiencies, and disadvantages, especially in the case of special educational needs, which prevent or hinder access to and permanence in the education system". The various regulations since then have taken up similar precepts of protection in line with the recommendations of specialised bodies (Gouédard et al., 2020; Pont & Montt, 2014; Schleicher, 2020) or the literature on educational return (Brunello & Paola, 2014; de la Fuente & Jimeno, 2009; Doncel et al., 2014) which establish the importance of the acquisition of competences, a fact which has become fundamental with the irruption of new technologies (Ballestar et al., 2020, 2022; Goos et al., 2009; Gregory et al., 2019).

Art.9 of the LOMLOE gives the responsibility to the autonomous communities (CCAA) to reduce and prevent school fail-

ure and early school leaving through territorial cooperation programmes. The Law's Analysis and Regulatory Impact Report (MAIN) in the first draft of the law, prior to the pandemic, provides 45 million euros per year for this type of programme for the period 2020-2023.

The autonomous regions were already investing heavily in this type of project. Since the 2007-2008 academic year, the Regional Ministry of Education of Castilla y León (Castile and León) has been developing, among others, the *Programme for the Improvement of Educational Achievement*. Its objectives are, in line with the LOE, LOMCE (Organic Law for the Improvement of the Quality of Education, 2013) and LOMLOE "... to contribute to the improvement of the educational model in the Community of Castilla y León and to facilitate the success and continuity of students in the education system" guaranteeing "the didactic progression of students and their promotion in the education system, reducing early school leaving, promoting their effective integration into the labour market and at the same time optimising the climate of coexistence in schools".

It is the schools that request these measures to favour educational success, the integration of students and the involvement of families to increase the promotion and graduation rate of students, with special attention to the most vulnerable students in the 4th year of ESO, 6th year of Primary Education and 2nd year of Baccalaureate in publicly funded schools. The programme was is taught in

93 centres in language, mathematics, and English to reinforce the end of the cycle and with the participation of students attending voluntarily and supported by their families. The access requirements are to require educational support or to be in a situation of socio-educational vulnerability and it is taught by specialised teachers.

For the 2019-2020 academic year, students participating in the programme increase their probability of passing language by 24.6%, English by 13.8% and mathematics by 8.4% compared to those not enrolled, generating a high level of satisfaction among both teachers (85.7%) and families (88.2%).

This first analysis serves as the basis for our research on the economic efficiency of the interventions included in the LOMLOE. To this end, we propose to classify students participating in the Programme for the Improvement of Educational Success in Castile and León in the 4th year of ESO into groups and to predict the probability of passing this school year for each of the groups of students. This will allow individual predictions to be made for students based on their characteristics, details of the special support programme carried out and their performance in the programme. This allows the probability of success of the programme to be estimated and its social return (ROI) calculated through a triangulation model that employs different ML methodologies such as CHAID decision trees and artificial neural networks to evaluate educational policies.

3. Empirical analysis.

The information used for the analysis contains the data available for students in the 4th year of ESO participating in the Programme for the Improvement of Educational Success in the 2019-2020 academic year. There are 1,739 records that correspond to the students who participated in one of the three programmes implemented in the Autonomous Community. Of these, 47.27% are girls and 52.73% are boys. Of these students, 47.15% (820 students) have already repeated a grade.

Students can participate in three categories of programmes. The C2 consists of accompanying students throughout the academic year and accounts for 76.37% of students (1,328). The C3 consists of extra classes during the summer in July and includes 17.02% of the students (296). The C2C3 programme, a combination of the two previous programmes, represents 6.61% of the students (115). These actions are delivered in the nine provinces of Castilla y León, with 66.36% of the students concentrated in provincial capitals and 33.64% in other localities, with a success rate in the completion of 4th-year ESO studies of 85.34% for those who finished the programme.

One of the innovations of our research is the application of a triangulation methodology, which consists of the development of more than one quantitative method, applying different approaches, with the aim of enriching the results, as well as confirming the results obtained twice. (Ballestar et al., 2020). In the first phase, we develop a Machine Learn-

ing (ML) model based on CHAID (Chi-square Automatic Interaction Detector) decision trees to determine which variables are relevant when designing and predicting the expected success rate of participants in the programme. The aim is to identify the factors to be considered when designing an intervention of these characteristics, as well as to calculate its success rate and, consequently, the social return on the investment made.

In the second phase, we assess the robustness of the first model by designing a new model using an ML methodology based on artificial neural networks multi-layer perceptron backward programming (ANN-MPL) to validate the results obtained previously.

3.1. Definition of variables in ML models.

CHAID decision trees and multi-layer perceptron artificial neural networks (ANN-MLP) are supervised machine learning methods. These models describe and explain the underlying relationships between different input variables in order to predict the value of the target variable, through a training process using a data sample that contains both the values for the input and output variables. (Maimon & Rokach, 2005). As supervised methods, the accuracy of their classifications and predictions will depend on the quality of the sample available for training, validation and testing of the models. (Aad et al., 2012; Ballestar et al., 2018; Li & Eastman, 2006).

We performed a data mining process on the database to create a single table for the development of the ML models. This table

contains 1,739 student records and 21 variables that capture the information for the characterisation of the student, as well as their performance and results both in the programme they have participated in and in the 4th year of ESO.

Through descriptive and iterative causal analyses, five of these variables were statistically significant and relevant for the em-

pirical analysis and development of the two ML models. Of these variables, four act as input variables in the models and one as an output or target variable to be determined or predicted (Table 1). Therefore, both models use the same five variables and are trained on the same database.

The target or output variable, the percentage of students who did not pass the

TABLE 1. Description of the variables of the ML models: CHAID decision tree and multilayer perceptron artificial neural network (ANN-MLP).

Input variables	Description
years_repetition	Discrete numeric variable. Number of years the student has repeated. If the student has not repeated any year, the value will be 0.
student_repetition	Boolean variable. Value 1 if the student has repeated a course; Value 0 if the student has not repeated a course.
kind_program	Categorical variable indicating which of the three academic programmes the student has completed: C2; C3; C2C3.
finish_program	Boolean variable. Value 1 if the student has completed the academic support programme; Value 0 if the student has not completed it.
Output variable	Description
finish_studies	Boolean variable. Value 1 if the student has passed the 4th year of ESO; Value 0 if the student has not passed the course.

Source: Own elaboration.

4th year of ESO (0 in the variable finish_studies), is 14.66%, while those who do pass the year represent 85.34% (value 1 of the variable finish_studies), leaving an unbalanced sample.

3.2. First phase of the empirical analysis: Predictive model of the success of the special education support programme.

In this phase, we developed a model consisting of a CHAID or *chi-squared automatic interaction detection* (CHAID) decision tree

(Kass, 1980)CHAID, is and offshoot of AID (Automatic Interaction Detection whose objective is twofold: classification and prediction. This model will classify each of the students into homogeneous groups based on the explanatory input variables or predictors that interact significantly with the dependent output or target variable (Akin et al., 2017; Khosravi et al., 2019). The algorithm identifies the students who are likely to belong to a particular group and, at the same time, identifies the rules that it



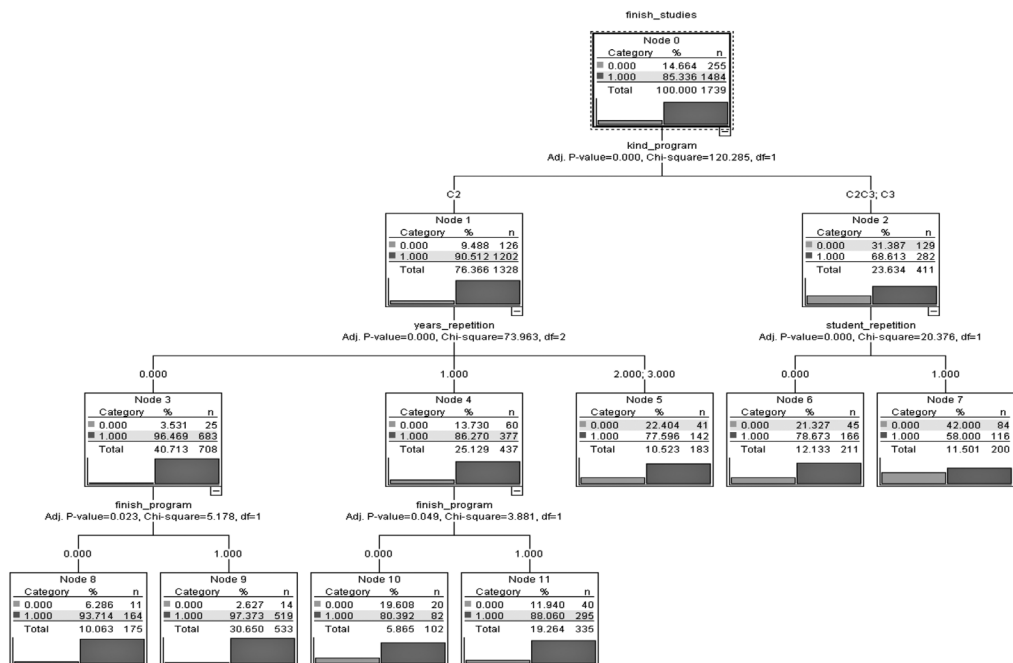
will use in predicting future cases of those students who have not participated. In our case, these groupings are constituted according to the characteristics of the student, the programme in which he/she participates and his/her performance, making it possible to fulfil the second objective of determining and predicting the pass rate in the 4th year of ESO for each of the groups and individually (Ramaswami & Bhaskaran, 2010). Finally, the prediction of this success rate makes it possible to calculate how to reinforce each of the programmes by type of student.

One of the benefits provided by CHAID decision trees over other methods is that they could generate non-binary trees. In addition, they also support continuous

or categorical numeric type input variables. Being non-parametric, it can handle both linear and non-linear relationships between the explanatory variables and the output variable and can also handle very large volumes of data very efficiently, even in real time, (Chassignol et al., 2018; Khosravi et al., 2019).

For the development and training of the model, we applied a training, testing and validation (TTV) methodology which consists of training the model with 70% (1,203 records) of the sample and testing and validating it with the remaining 20% (367 records) and 10% (169 records) of the sample, respectively. (Ballestar et al., 2019) because of the difficulty of identifying who are the main beneficiaries and

GRAPH 1. CHAID decision tree diagram for the success model of the student support programmes for the successful completion of the 4th year of ESO.



Source: Own elaboration.

what are the long-term effects. Still, new policies including financial incentives have been adopted to increase the research output at all possible levels. Little literature has been devoted to the response to those incentives. To bridge this gap, we carry out our analysis with data of a six years program developed in Madrid (Spain). Graph 1 shows the CHAID decision tree diagram obtained for the programme success model.

3.2.1. First phase of the empirical analysis: Evaluation of the model.

The most relevant CHAID indicators are classification accuracy, sensitivity, specificity, area under the ROC curve and the GINI coefficient. These accuracy indicators and the confusion matrix, which contains the percentage of cases classified both correctly and incorrectly for the two possible values of the dependent variable, are available in Table 2. The indicators have been calculated both for the total sample and for each of the training, testing and validation subsamples of the model. Therefore, we observe that the model has an accuracy in classifying students and predicting their percentage of passing students of 70.73% (error of 29.27%). This percentage of correctly classified students is very similar between all the training, testing and validation subsamples, so we can affirm that the model has not been overtrained.

The percentage of true positives, called *sensitivity*, is 71.43%. This value explains the percentage of students who pass the course and who have been correctly classified by the CHAID decision tree based

on the characteristics of the student, the support programme carried out and their result in this programme. The percentage of true negatives, also called *specificity*, is 66.67%. This value is the percentage of students who do not pass the 4th year of ESO and who have been correctly classified based on the same input variables. The complementary values are the percentage of false positives, 33.33%, which corresponds to the percentage of students who did not pass the 4th year of ESO and were classified by the model as having achieved it. Finally, the percentage of false negatives is 28.57%. This value corresponds to the percentage of students who, having passed the year, were classified by the model as not having passed it.

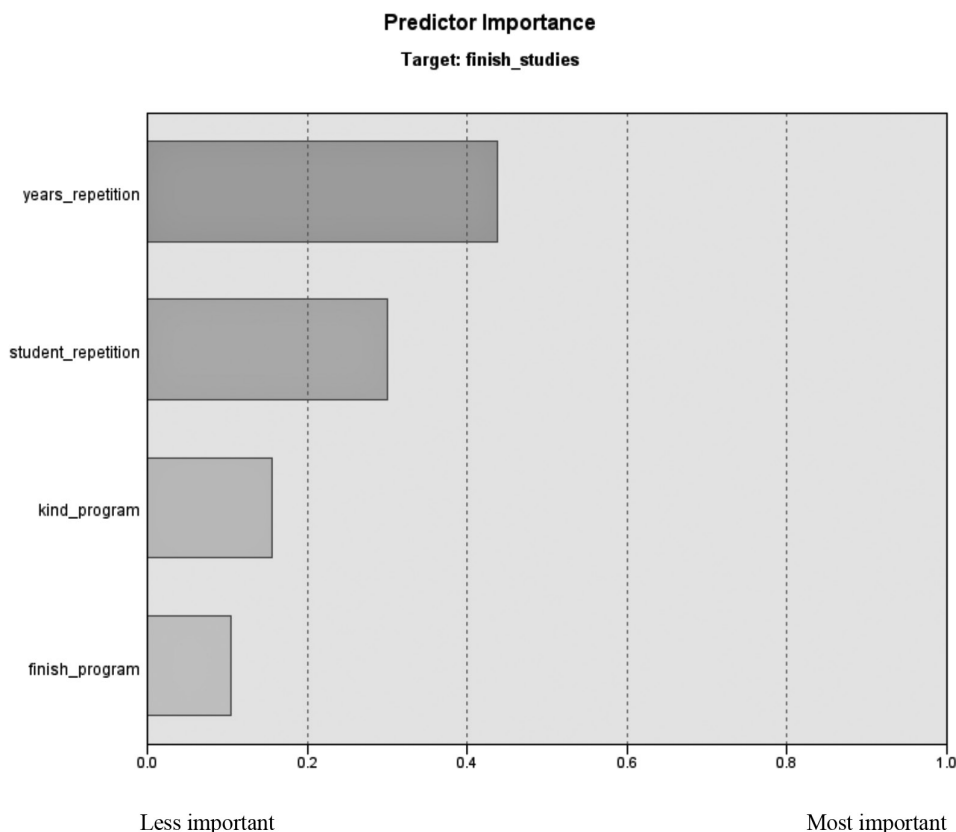
In this research we prefer to use as the main measure of accuracy of the CHAID decision tree the area under the curve (AUC) ROC indicator, as it is more robust than the classification accuracy indicator when working with unbalanced samples as in our case (Table 2). (Dželihodži & Jonko, 2016; Yin et al., 2013). The area under the ROC curve (AUC) for the total sample has a value of 0.762, like those of the training, testing and model validation subsamples (0.763, 0.760 and 0.755 respectively). The AUC values can range from 0.5, which is the worst possible value and would imply that the model makes random classifications, to 1, which is the best value and would mean that the model makes perfect classifications. Therefore, it can be concluded that the quality of this CHAID decision tree model is good. (Hosmer Jr., et al., 2013). Complementarily, the GINI coefficient has also been calculated.

TABLE 2. Model accuracy and confusion matrix of the CHAID decision tree.

Model Accuracy		Confusion Matrix				Predicción				
						Sample size		Percentage		
Sample	Sample size	Percentage Correct	AUC	GINI	Sample	Observed	0	1	0	1
Total Sample	1739	70.73%	0.762	0.523	Total Sample	0	170	85	66.67%	33.33%
						1	424	1060	28.57%	71.43%
Sub-sample										
Training	1203	70.91%	0.763	0.527	Training	0	117	62	65.36%	34.64%
						1	288	736	28.13%	71.88%
Test	367	69.21%	0.760	0.519	Test	0	34	16	68.00%	32.00%
						1	97	220	30.60%	69.40%
Validation	169	72.78%	0.755	0.510	Validation	0	19	7	73.08%	26.92%
						1	39	104	27.27%	72.73%

Source: Own elaboration.

GRAPH 2. Relative importance of input variables in the CHAID decision tree.



Source: Own elaboration.

ed, which can be used alternatively to the AUC, as they are closely related. The GINI coefficient represents twice the area between the ROC curve and the diagonal and ranges between the values 0 and 1. In our research, the GINI coefficient for the total sample has a value of 0.523, like that of the training, testing and validation subsamples of the model (0.527, 0.519 and 0.510 respectively).

Graph 2 shows the relative importance of each of the significant predictor variables in the CHAID decision tree.

The importance of these predictors is determined by calculating the reduction in variance of the target variable (completion or non-completion of the 4th year of ESO per student) that can be attributed to each predictor by performing a sensitivity analysis. (Saltelli et al., 2004). In this research, the variable that accumulates 44% of the relative importance is the number of years that the student has repeated (years_repetition), followed by the variable indicating whether the student has repeated or not (student_repetition) with an importance of 30%. Next, with

16%, is the type of programme in which the student participated (*kind_program*) and, finally, with 10%, whether the student completed the special support programme (*finish_program*).

3.2.2. First phase of the empirical analysis: Analysis of the results of the CHAID decision tree model.

The CHAID-based ML model produces a tree in which the first group, called the *root node*, represents the total set of the sample, in this case, the 1,739 students analysed. The algorithm divides this set into two or more categories called *parent or initial nodes*. Below the parent nodes, child nodes are linked to them. The categories at the last level of the decision tree are called *terminal nodes*. In terms of hierarchy, the *parent or initial nodes* exert the greatest influence on the *root node*, corresponding to the dependent variable it seeks to explain, while they exert a lesser influence on the *terminal nodes*.

The total number of nodes is 12, distributed as follows in three levels of depth: one root node (level 0), four parent nodes (distributed in levels 1 and 2) and seven terminal nodes (distributed in levels 2 and 3) (Graph 1). Additionally, based on the model evaluation analysis carried out in the previous section, it has been shown that this model has a good segmentation and prediction capacity (ROC of 0.762).

3.2.2.1. Level 0 of the CHAID Decision Tree.

At Level 0 of the CHAID decision tree, Node 0 is the variable to be predicted, which determines whether the student

has successfully completed the 4th year of ESO (*finish_studies*). At this level it is observed that, for the total sample, 85.34% have passed the course, while 14.66% have not.

3.2.2.2. Level 1 of the CHAID Decision Tree.

At this level, the most relevant variable to define the first two parent nodes (Node 1 and Node 2) is the type of programme in which the student has participated (*kind_program*). Each of these nodes will give rise to a sub-tree that will use different predictor variables to define the following levels. At this level, the most relevant variable is the programme in which the student has participated, whether he/she has completed it or not. Of the students who have participated in the C2 programme (accompaniment of the student throughout the 4th year of ESO), 90.51% have passed the 4th year of ESO, compared to 68.61% of those who have participated in the C3 and C2C3 programmes. Therefore, the programme with the greatest success in its objectives is C2.

3.2.2.3. Level 2 of CHAID Decision Sub-tree 1 (Left).

The next most relevant variable having participated in the C2 programme is the number of times they have repeated a year (*years_repetition*). In Node 1 90.51% of the students pass, but this percentage can vary greatly depending on whether the student has repeated a year or not previously: students who have not repeated a year increase their pass rate to 96.46% (Node 3), while those who have repeated once see it reduced to 86.27% (Node 4) and those who have repeated two or more times to 77.59% (Node

5). Non-repeaters obtain better results in the 4th year of ESO compared to those who have repeated a year at least once.

In this Level 2 we find the nodes Node 3, 4 and 5. On the one hand, the nodes Node 3 and 4, in which the students have not repeated a grade or have done so only once, give rise to the new subtrees of Level 3. In this Level 3 the most discriminant variable is the one that determines whether the student managed to finish the C2 programme in which he/she participated (finish_program). On the other hand, Node 5 of Level 2 is a terminal node (it has no more sublevels), which implies that the probability of passing the 4th year of ESO for the students in this group is independent of their performance in the C2 programme, contrary to Node 3 and 4.

3.2.2.4. Level 3 of CHAID Decision Subtree 1 (Left).

In Level 3 there are 4 terminal nodes: Node 8 and Node 9 correspond to students who, having participated in C2, have not repeated previously, and come from the sub-tree of Node 3. The students in Node 9 have completed the C2 programme and therefore 97.37% have managed to pass the 4th year of ESO, the most successful figure of all the segments of the CHAID decision tree. While in Node 8 they have not managed to complete the C2 programme, and their success rate drops to 93.71%.

Node 10 and Node 11 correspond to students in C2 who have repeated a grade once and come from the sub-tree of Node 4. Students in Node 11 have completed the C2 programme and therefore 88.06% have

managed to pass the 4th year of ESO, while those in Node 10 see this figure reduced to 80.39% because they have not completed the C2 programme. Therefore, students who complete the C2 programme obtain better results in the 4th year of ESO compared to those who have not completed the programme.

3.2.2.5. Level 2 of CHAID Decision Subtree 2 (Right).

In level 2 of subtree 2 (right, having been part of C2C3 or C3) is whether the student has ever repeated a year (student_repetition). We know that in Node 2 only 68.61% of the students pass, but this percentage can vary a lot depending on whether the student is a repeater or not. Students who have not repeated any year increase their pass rate in 4th ESO to 78.67% (Node 6), while those who have repeated a year see this percentage reduced to 58% (Node 7). Therefore, non-repeaters obtain better results in the 4th year of ESO compared to those who have repeated a grade. In this sub-tree 2 (right) there are no more levels, and this variable is the last relevant variable to generate more groups of students.

3.3. Second phase of the empirical analysis: Analysis of the robustness and predictive quality of the model.

In this research, a triangulation methodology is applied to test the robustness of the results obtained in the first phase. For this purpose, an ML predictive model based on a multilayer perceptron artificial neural network (ANN-MLP) is developed. Both models use the same database with the same input variables (student

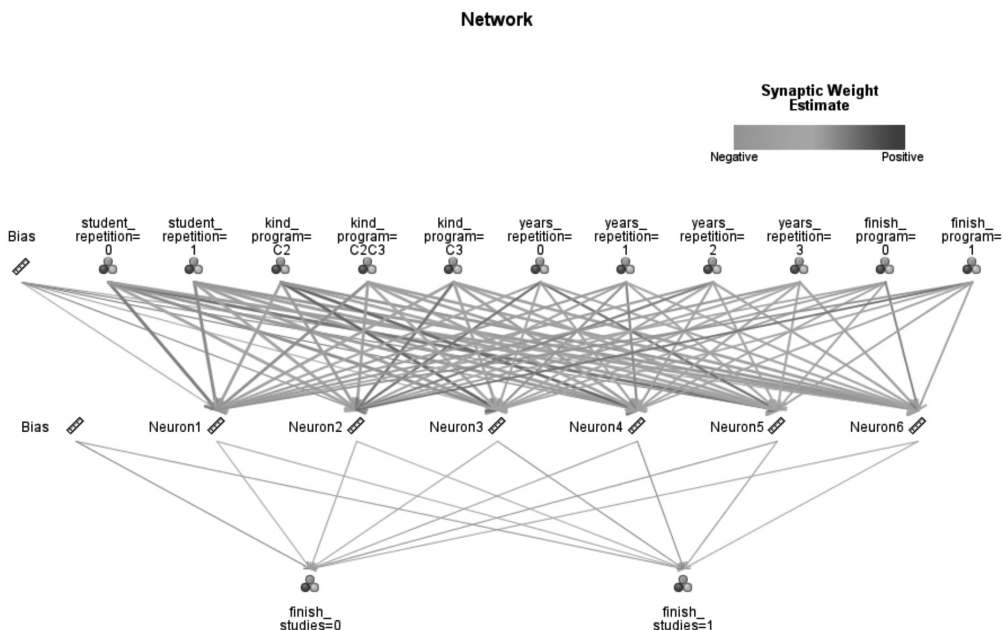
characterisation, performance and type of programme carried out) and output (student performance in the 4th year of ESO) (Table 1), but they are built using completely different ML methods, with the aim of verifying whether they obtain confirmatory results. (Wolszczak-Derlacz & Parteka, 2011).

ANN-MLP has the ability of neural networks to handle complex relationships between variables, both linear and non-linear, which makes them an alternative to more traditional predictive models such as logistic regressions where the variable to be predicted is dichotomous. (Paliwal &

Kumar, 2009). This ensures that no relationships are left unidentified.

While training the ANN-MLP it should be considered whether the distribution of the dependent variable is unbalanced. This bias in our sample could have a negative impact on the ANN-MLP training process. To avoid this, we applied an *oversampling method* on the underrepresented group to train the model with 70% (2,098 records) of the sample already balanced, ensuring that both groups of students are equally represented when training the model.

GRAPH 3. ANN-MLP architecture for predicting the success of programmes to support students to pass the 4th year of ESO.



Less important

Most important

Source: Own elaboration.

On the other hand, the validation and testing of the ANN-MLP model is carried out respectively with 20% (367 records) and 10% (169 records) of the sample, this time, without balancing. We also follow a training, testing and validation (TTV) methodology.

Graph 3 shows the architecture used to predict the success of student support programmes for passing the 4th grade of ESO, consisting of three layers: the first input layer consists of eleven units, corresponding to the possible values of the four input variables related to the characteristics of the students, their performance and the programme carried out (Table 1). The middle layer, called the hidden layer, consists of six units or neurons. Finally, the last one consists of two units, corresponding to the prediction of student performance in 4th ESO (Table 1). The type of activation function of the hidden layer and the output layer correspond to a hyperbolic tangent and *softmax* respectively.

Table 3 describes the accuracy indicators and the confusion matrix of the ANN-MLP which show that the quality of this model is good: correct classification of 71.19% of the students and an AUC of 0.763. (Hosmer Jr., et al., 2013). Furthermore, it is also verified that there is no over-training as the indicators obtained for the total sample, as well as for the training, testing and validation subsamples are similar.

4. Discussion

The indicators we have used to assess the accuracy of the ANN-MLP as a pre-

dictor of student success are the same as those used for the CHAID, thus facilitating the comparison between results, predictive quality, and robustness of both models.

Table 4 compares the classification accuracy and confusion matrix indicators of both models (Table 3 vs. Table 2). The predictions and accuracy of the ANN-MLP are consistent with those of the CHAID decision tree. The ANN-MLP correctly predicts 71.19% (8.81% error) of the students who will pass the 4th year of ESO. This figure represents a 0.46% (8 students) higher predictive capacity compared to the 70.73% of correct predictions obtained by the CHAID decision tree. For the rest of the indicators, such as sensitivity, specificity, false positives and false negatives, the differences are also statistically non-significant, ranging from -1.57% (difference in the classification of four students) to 0.81% (difference in the classification of 12 students). The same is true for the differences between the areas under the ROC curve (AUC) and the GINI coefficient. The AUC for the ANN-MLP and the CHAID decision tree, 0.763 and 0.762 respectively and the GINI coefficient 0.525 and 0.523 respectively. Graph 4 shows the relative importance of the ANN-MLP input variables or predictors following the variance method and in total they add up to 1.

Although it is true that both the CHAID decision tree and the ANN-MLP use the same database and variables in the model and manage to achieve very similar results, predictive capacity and robustness, the way of doing so is very different, both in the methods used

TABLE 3. Model accuracy and confusion matrix of the artificial neural network multilayer perceptron (ANN-MLP).

Model Accuracy			Confusion Matrix							
			Prediction							
Sample	Sample Size	Percentage Correct	AUC	GINI	Sample	Observed	Sample Size		Percentage	
Total Sample	1739	71.19%	0.763	0.525	Total Muestra	0	166	89	65.10%	34.90%
						1	412	1072	27.76%	72.24%
Submuestra										
Training	2098*	69.78%	0.767	0.535	Training	0	720	354	67.04%	32.96%
						1	280	744	27.34%	72.66%
Test	367	70.03%	0.755	0.510	Test	0	31	19	62.00%	38.00%
						1	91	226	28.71%	71.29%
Validation	169	69.23%	0.744	0.488	Validation	0	15	11	57.69%	42.31%
						1	41	102	28.67%	71.33%

*Oversampled to balance the sample.

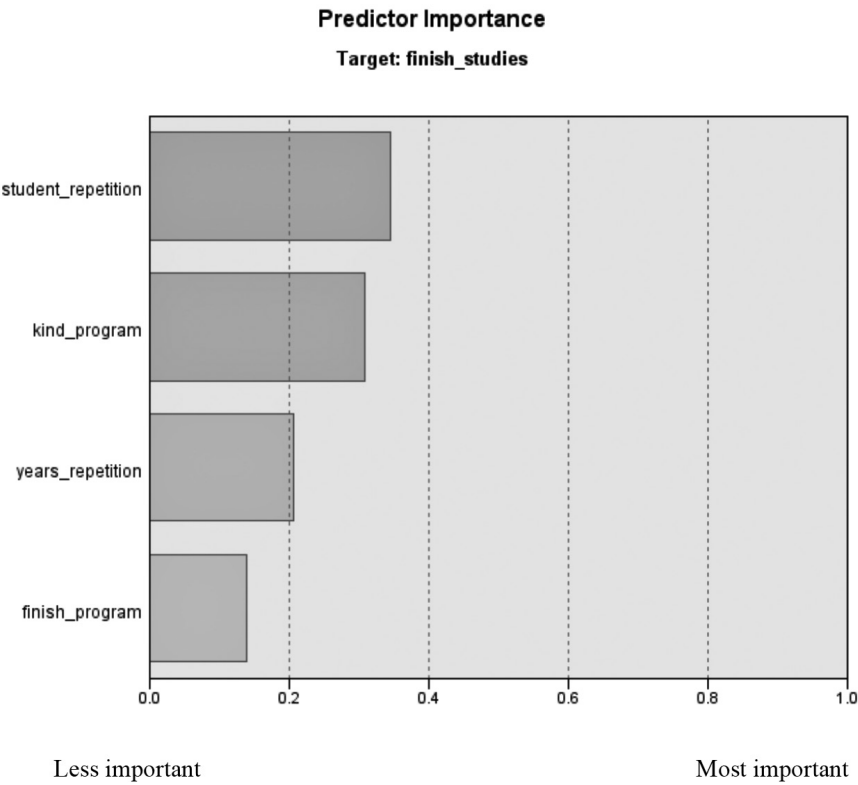
Source: Own elaboration.

TABLE 4. Comparison of indicators: Classification accuracy and confusion matrix of the ANN-MLP compared to the CHAID decision tree.

Indicators	ANN-MLP compared to CHAID Tree	
	% change	Classified variation
Percentage Correct	0.46%	8
Sensitivity	0.81%	12
Specificity	-1.57%	-4
False positives	1.57%	4
False negatives	-0.81%	-12
AUC	0.001	-
GINI	0.002	-

Source: Own elaboration.

GRAPH 4. Relative importance of the input variables of the artificial neural network multilayer perceptron (ANN-MLP).



Source: Own elaboration.

and, in the handling, and importance of the variables that contribute to the model. In the case of the ANN-MLP, the variable with the highest relative importance accumulates up to 35% and indicates whether the student has repeated a year (student_repetition), followed by the type of programme in which the student participated with an importance of 31% (kind_program), next with an importance of 21% is the variable indicating the number of years repeated by the student (years_repetition) and, finally, with an importance of 13% is the variable indicating whether the student finished the special support programme (finish_program).

The application of the triangulation methodology concludes that the CHAID decision tree models and the ANN-MLP obtain similar results in terms of classification accuracy and prediction of the success of the student support programmes for passing the 4th year of ESO and, therefore, they are robust models that confirm their results. Of the three programmes analysed, the one with the highest success rate is C2, which accompanies the student throughout the entire 4th year of ESO, reaching a 90.51% pass rate, compared to 68.61% for the other two options C2 (extraordinary classes in July) and C2C3 (combination of the previous options).

In terms of educational evaluation, the results are striking: the economic and social return of an intervention can be evaluated according to the different characteristics of the students, such as course performance, the type of course or, of course, whether they are repeaters or not.

5. Conclusions

Through our research we have contrasted the need for an economic evaluation of the LOMLOE. The law started with the same budget allocation as the LOMCE for interventions such as the one analysed (45 million euros). The arrival of European funds opens up new possibilities not contemplated in its initial design, which in turn raises the challenge of using this additional investment efficiently.

Our analysis shows that it is possible to predict the probability of at-risk students passing the 4th year of ESO depending on the support programme they participate in. This can contribute to the design of more efficient programmes with higher success rates.

To complement the models already used by Public Administrations, we employ robust ML models such as CHAID decision trees and artificial neural networks to analyse the characteristics of the students, the intervention they have taken part in and the final outcomes. This allows us to maximise the social return of each programme and to support those students who are less likely to succeed.

For example, our results show that, in the *Programme for the Improvement of Educational Success* in Castile and Leon, regardless of the programme in which the student has participated, those who have never repeated a grade are more likely to pass the 4th year of ESO than those who have. The probability of passing the year increases when students, in addition to participating in special programmes, stay for the full support program. For this rea-

son, the group of students who achieve the best results is the group of non-repeaters who participate in and complete the C2 programme, with 97.37% passing the 4th year of ESO. While the group with the worst results is the group of repeaters who have participated in the C2C3 and C3 programmes, regardless of whether they have completed them or not, with a 58% pass rate in the 4th year of ESO. Our analysis also shows that providing educational support in July does not constitute a great added value since the C1 programme with accompaniment during the school year obtains better results than the C3 programme which combines this reinforcement during the school year with an additional reinforcement in the month of July. In the next phase of research, we will check whether these different results are due to the pupils themselves or to the characteristics of the programme.

As a corollary, we also show that ML models such as CHAID and artificial neural networks are good candidates for building models to answer the educational questions posed, such as the importance of mentoring the student to complete the programme and not drop out, especially in the case of young people who have repeated, for whom the programme is a very efficient tool to avoid school failure and an efficient way of investing public money.

Note

1 <https://educationendowmentfoundation.org.uk/projects-and-evaluation/projects/the-impact-of-covid-19-on-school-starters/>

References

- Aad, G., Abbott, B., Abdallah, J., Abdelalim, A. A., Abdesselam, A., Abidinov, O., Abi, B., Abolins, M., Abramowicz, H., Abreu, H., Acerbi, E., Acharya, B. S., Adams, D. L., Addy, T. N., Adelman, J., Aderholz, M., Adomeit, S., Adragna, P., Adye, T., ... Zwalinski, L. (2012). Performance of missing transverse momentum reconstruction in proton-proton collisions at $\sqrt{s} = 7$ TeV with atlas. *European Physical Journal C*, 72 (1), 1-35. <https://doi.org/10.1140/epjc/s10052-011-1844-6>
- Akin, M., Eydurán, E., & Reed, B. M. (2017). Use of RSM and CHAID data mining algorithm for predicting mineral nutrition of hazelnut. *Plant Cell, Tissue and Organ Culture (PCTOC)*, 128 (2), 303-316.
- Ballestar, M. T., Doncel, L. M., Sainz, J., & Ortigosa-Blanch, A. (2019). A novel machine learning approach for evaluation of public policies: An application in relation to the performance of university researchers. *Technological Forecasting and Social Change*, 149. <https://doi.org/10.1016/j.techfore.2019.119756>
- zBallestar, M. T., García-Lazaro, A., Sainz, J., & Sanz, I. (2022). Why is your company not robotic? The technology and human capital needed by firms to become robotic. *Journal of Business Research*, 142, 328-343. <https://doi.org/10.1016/j.jbusres.2021.12.061>
- Ballestar, M. T., Grau-Carles, P., & Sainz, J. (2018). Customer segmentation in e-commerce: Applications to the cashback business model. *Journal of Business Research*, 88, 407-414. <https://doi.org/10.1016/j.jbusres.2017.11.047>
- Brunello, G., & Paola, M. De. (2014). The costs of early school leaving in Europe. *IZA Journal of Labor Policy*, 3 (1), 1-31. <https://doi.org/10.1186/2193-9004-3-22In>
- Burgess, S. (2020). *How we should deal with the lockdown learning loss in England's schools*. VOX Cepr Policy Portal. <https://bit.ly/3Eom12M>
- Chassignol, M., Khoroshavin, A., Klimova, A., & Bilyatdinova, A. (2018). Artificial Intelligence trends in education: A narrative overview. *Procedia Computer Science*, 136, 16-24. <https://doi.org/10.1016/j.procs.2018.08.233>

- Chatterjee, S., & Bhattacharjee, K. K. (2020). Adoption of artificial intelligence in higher education: a quantitative analysis using structural equation modelling. *Education and Information Technologies*, 25 (5), 3443-3463. <https://doi.org/10.1007/s10639-020-10159-7>
- Crescenzi, R., Giua, M., & Sonzogno, G. V. (2021). Mind the Covid-19 crisis: An evidence-based implementation of Next Generation EU. *Journal of Policy Modeling*, 43 (2), 278-297. <https://doi.org/10.1016/J.JPOLMOD.2021.03.002>
- De la Fuente, A., & Jimeno, J. F. (2009). The private and fiscal returns to schooling in the European Union. *Journal of the European Economic Association*, 7 (6), 1319-1360. <https://doi.org/10.1162/JEEA.2009.7.6.1319>
- Doncel, L. M., Durá, P., Grau, P., & Sainz, J. (2014). Fiscal sustainability and immigration in the Madrid Region. *International Migration*, 52 (6), 180-196. <https://doi.org/10.1111/imig.12004>
- Dželihođić, A., & Jonko, D. (2016). Comparison of ensemble classification techniques and single classifiers performance for customer credit assessment. *Modeling of Artificial Intelligence*, 3, 140-150.
- European Commission (2010). *EUROPE 2020: A strategy for smart, sustainable and inclusive growth*. <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2010:2020:-FIN:EN:PDF>
- Golden, G. (2020). *OECD iLibrary | Education policy evaluation: Surveying the OECD landscape*. (OECD Education Working Papers, 236). OECD Publishing. https://www.oecd-ilibrary.org/education/education-policy-evaluation_9f127490-en
- Goos, M., Manning, A., & Salomons, A. (2009). Job polarization in Europe. *American Economic Review*, 99 (2), 58-63. <https://doi.org/10.1257/aer.99.2.58>
- Gouëdard, P., Pont, B., & Viennet, R. (2020). Education responses to COVID-19: Implementing a way forward. *OECD Education Working Papers*, 224. OECD Publishing. <https://doi.org/10.1787/8e95f977-en>
- Gregory, T., Salomons, A., & Zierahn, U. (2019). Racing With or Against the Machine? Evidence from Europe. *IZA Institute of Labor Economics*, 2019 (12063).
- Hosmer Jr., D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression*. John Wiley & Sons.
- Kass, G. V. (1980). An exploratory technique for investigating large quantities of categorical data. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 29 (2), 119-127. <https://doi.org/10.2307/2986296>
- Khosravi, P., Kazemi, E., Zhan, Q., Malmsten, J. E., Toschi, M., Zisimopoulos, P., Sigaras, A., Lavery, S., Cooper, L. A. D., & Hickman, C. (2019). Deep learning enables robust assessment and selection of human blastocysts after in vitro fertilization. *NPJ Digital Medicine*, 2 (1), 1-9.
- Kraft, M. (2015, August 20). What If Every Struggling Student Had a Tutor? *Education Week*. <https://www.edweek.org/education/opinion-what-if-every-struggling-student-had-a-tutor/2015/08>
- Li, Z., & Eastman, J. R. (2006). Commitment and typicality measurements for fuzzy ARTMAP neural network. *Geoinformatics 2006: Geospatial Information Science*, 6420. <https://doi.org/10.1117/12.7129986420>, 642011.
- Maimon, O., & Rokach, L. (2005). *Data mining and knowledge discovery handbook*. Springer.
- Nickow, A., Oreopoulos, P., & Quan, V. (2020). The impressive effects of tutoring on prek-12 learning: A systematic review and meta-analysis of the experimental evidence. *National Bureau of Economic Research Working Papers Series*, 27476. <https://doi.org/10.3386/w27476>
- OECD. (2018). *Education policy outlook 2018: Putting student learning at the centre*. OECD.
- Organic Law 10/2002, of 23 December 2002, on the Quality of Education (LOCE). *Spanish Official State Gazette*, 307, of 24 December 2002, pp. 45188 to 45220. <https://www.boe.es/eli/es/lo/2002/12/23/10>
- Organic Law 2/2006, of 3 May, on Education (LOE). *Spanish Official State Gazette*, 106. <https://www.boe.es/eli/es/lo/2006/05/03/2/con>
- Organic Law 8/2013, of 9 December, for the Improvement of the Quality of Education (LOMCE). *Spanish Official State Gazette*, 295, pages 97858 to 97921. <https://www.boe.es/eli/es/lo/2013/12/09/8>

Organic Law 3/2020, of 29 December, which amends Organic Law 2/2006, of 3 May, on Education (LOMLOE). *Spanish Official State Gazette*, 340, 30 December 2020, pages 122868 to 122888. https://www.boe.es/diario_boe/txt.php?id=BOE-A-2020-17264

Organic Law 3/2020, of 29 December, which amends Organic Law 2/2006, of 3 May, on Education (LOMLOE). *Spanish Official State Gazette*, 340, 30 December 2020, pages 122868 to 122888. https://www.boe.es/diario_boe/txt.php?id=BOE-A-2020-17264

Paliwal, M., & Kumar, U. A. (2009). Neural networks and statistical techniques: A review of applications. *Expert Systems with Applications*, 36 (1), 2-17.

Pont, B., & Montt, G. (2014). *Education policy outlook. Spain*. OECD.

Porte, C. de la, & Jensen, M. D. (2021). The next generation EU: An analysis of the dimensions of conflict behind the deal. *Social Policy & Administration*, 55 (2), 388-402. <https://doi.org/10.1111/SPOL.12709>

Ramaswami, M., & Bhaskaran, R. (2010). A CHAID based performance prediction model in educational data mining. *International Journal of Computer Science*, 7 (1), 10-18.

Saltelli, A., Tarantola, S., Campolongo, F., & Ratto, M. (2004). *Sensitivity analysis in practice: a guide to assessing scientific models (Vol. 1)*. Wiley Online Library.

Schleicher, A. (2020). The impact of COVID-19 on education: Insights from education at a glance 2020. *OECD Journal: Economic Studies*. <https://www.oecd.org/education/the-impact-of-covid-19-on-education-insights-education-at-a-glance-2020.pdf>

Slavin, R. E. (2016). Evidence-Based Education Policies: Transforming Educational Practice and Research. *Educational Researcher*, 31 (7), 15-21. <https://doi.org/10.3102/0013189X031007015>

Wolszczak-Derlacz, J., & Parteka, A. (2011). Efficiency of European public higher education institutions: A two-stage multicountry approach. *Scientometrics*, 89 (3), 887-917. <https://doi.org/10.1007/s11192-011-0484-9>

Yin, L., Ge, Y., Xiao, K., Wang, X., & Quan, X. (2013). Feature selection for high-dimensional imbalanced data. *Neurocomputing*, 105, 3-11.

Yusuf, A. (2007). *Standard-Setting at UNESCO: Conventions, recommendations, declarations and charters adopted by UNESCO, 1948-2006 (Vol. 2)*. Martinus Nijhoff Publishers.

Authors' biographies

María Teresa Ballestar holds a degree in Statistics from the University of Zaragoza, a degree in Market Research and Techniques and a Master's degree in Information and Knowledge Society from the Universitat Oberta de Catalunya. PhD in Big Data methodologies and technologies applied to Economics at the Universidad Rey Juan Carlos. She is also an Associate Professor at the Universidad Rey Juan Carlos. In recent years she has held management positions and led projects in digital transformation, innovation, data analytics and data science. She has published more than a dozen scientific papers and collaborated with various media.



<https://orcid.org/0000-0001-8526-7561>

Jorge Sainz Orcid holds a degree in Economics from the Universidad Complutense de Madrid and in Law from the UNED. He holds a PhD in Economics (URJC) and an MBA (specialising in finance and public policy) from the Simon School, University of Rochester. He is Professor of Applied Economics at the Universidad Rey Juan Carlos de Madrid and Visiting Fellow at the Institute for Policy Research, University of Bath.

He has been Advisor to the Cabinet of the Regional Minister of Education and General Deputy Director of Research at the Regional Ministry of Education of the Community of Madrid. In the Ministry of



Education, Culture and Sport he has been General Director of University Policy and General Secretary of Universities.



<https://orcid.org/0000-0001-8491-3154>

Ismael Sanz is Associate Professor in the Department of Applied Economics I at the Universidad Rey Juan Carlos. He holds a PhD in Applied Economics from the Faculty of Economics and Business Studies at the Universidad Complutense de Madrid (UCM). He has been General

Director of Innovation, Scholarships and Grants of the Ministry of Education of the Community of Madrid, Director of the National Institute for Educational Evaluation (INEE) of the Ministry of Education, Culture and Sport (2012-2015) and Chair of the Strategic Development Group of PISA of the OECD. He is currently Vice-Rector for Quality at the URJC. Author of articles in high impact research journals.



<https://orcid.org/0000-0003-1286-4124>

Table of contents

Sumario

The LOMLOE amidst the challenges of the Spanish education system in the 21st century ***La LOMLOE ante los desafíos de la educación española en el siglo XXI***

Guest editor: Francisco López Rupérez

Editor invitado: Francisco López Rupérez

Francisco López Rupérez

Presentation: The LOMLOE amidst the challenges of the Spanish education system in the 21st century

Presentación: La LOMLOE ante los desafíos de la educación española en el siglo XXI

3

Studies and essays ***Estudios y ensayos***

Francisco Esteban Bara, & Fernando Gil Cantero

The purposes of education and Spain's LOMLOE educational legislation: Controversial questions in educational action

Las finalidades de la educación y la LOMLOE: cuestiones controvertidas en la acción educativa

13

José Luis Gaviria, & David Reyero

The transmission of cultural content and its evaluation among the ends of the education system: An analysis of the LOMLOE

La transmisión de los contenidos culturales y su evaluación entre los fines del sistema educativo, según la LOMLOE

31

Francisco López Rupérez

The competency-based curriculum approach. An analysis of the LOMLOE

El enfoque del currículo por competencias.

Un análisis de la LOMLOE

55

José Luis Martínez López-Muñiz

The regulatory quality of legal frameworks:

A critical approach

La calidad regulatoria de los marcos normativos: una aproximación crítica

69

Charles L. Glenn

Educational pluralism and vulnerable children

Pluralismo educativo y niños vulnerables

85

José Adolfo de Azcárraga

The new Spanish educational legislation:

Why public education will not improve

La nueva legislación educativa: por qué no mejorará la educación pública en España

111

Testing and proposals *Investigaciones y propuestas*

María Teresa Ballestar, Jorge Sainz, & Ismael Sanz

An economic evaluation of educational interventions in the LOMLOE: Proposals for improvement with Artificial Intelligence
Evaluación económica de intervenciones educativas en la LOMLOE: propuestas de mejora con inteligencia artificial 133

Francisco López Rupérez

The quality of governance of the education system.
The case of the LOMLOE
La calidad de la gobernanza del sistema educativo. El caso de la LOMLOE 155

Inmaculada Egido Gálvez

Curriculum reform to respond to the challenges of the future. Spain in international perspective
La reforma del currículo para responder a los retos del futuro. España en perspectiva internacional 175

Samuel Gento Palacios, Raúl González-Fernández, & Ernesto López-Gómez

Heads of educational institutions and expansion of autonomy with accountability. The mediating role of pedagogical leadership
Dirección de centros educativos y ampliación de autonomía con rendición de cuentas. El rol mediador del liderazgo pedagógico 193

Book reviews *Reseñas bibliográficas*

Escámez-Sánchez, J., & Peris-Cancio J.-A. (2021).

La universidad del siglo XXI y la sostenibilidad social [The 21st century university and social sustainability] (Ramón Mínguez-Vallejos). **Grupo SI(e)TE Educación (2021).**

La calidad en la educación [Quality in education] (José Antonio Jordán). **Santos-Rego, M. A., Lorenzo, M., & Mella, I. (2020).** *El aprendizaje-servicio y la educación universitaria. Hacer personas competentes [Service-learning and university education. Making competent people]* (Alexandre Sotelino Losada). 211

This is the English version of the research articles and book reviews published originally in the Spanish printed version of issue 281 of the **revista española de pedagogía**. The full Spanish version of this issue can also be found on the journal's website <http://revistadepedagogia.org>.



ISSN: 0034-9461 (Print), 2174-0909 (Online)

<https://revistadepedagogia.org/>

Depósito legal: M. 6.020 - 1958

INDUSTRIA GRÁFICA ANZOS, S.L. Fuenlabrada - Madrid