

# Contribución del *machine learning* al análisis de la repetición escolar en España: un estudio con datos PISA

## Contribution of machine learning to the analysis of grade repetition in Spain: A study based on PISA data

**Alexander CONSTANTE-AMORES.** Profesor Asociado. Universidad Camilo José Cela, España (*israelalexander.constante@ucjc.es*).

**Dra. Delia ARROYO-RESINO.** Profesora Contrada Doctora. Universidad Complutense de Madrid, España (*garroy01@ucm.es*).

**María SÁNCHEZ-MUNILLA.** Investigadora en Formación. Universidad Complutense de Madrid, España (*marsan27@ucm.es*).

**Dra. Inmaculada ASENSIO-MUÑOZ.** Profesora Titular. Universidad Complutense de Madrid, España (*macu@edu.ucm.es*).

### Resumen:

**Introducción:** La repetición de curso tiene una excesiva incidencia en España a pesar de ser una medida controvertida. A fin de obtener evidencias que contribuyan a su reducción en la educación obligatoria, el presente trabajo profundiza en el estudio de los índices de contexto de PISA 2018 más vinculados con dicho fenómeno. **Método:** Con la muestra de estudiantes españoles ( $n = 35\,943$ ), se utiliza un método de aprendizaje automático para seleccionar y ordenar los predictores, y una regresión logística multinivel (estudiantes y centros) para cuantificar la contribución de cada uno. **Resultados:** Se obtienen las 30 va-

riables de contexto más relevantes en cada etapa educativa, que explican el 65.5% de la varianza de la repetición en primaria y casi el 55.7% en secundaria. **Conclusiones:** Los principales indicadores son sobre todo del nivel de estudiantes, lo que sugiere la idoneidad de intervenciones psicoeducativas basadas en el apoyo individualizado más que en políticas generalizadas. De ahí emergen medidas potencialmente más eficientes y equitativas que la repetición, centradas, por ejemplo, en la gestión del tiempo de aprendizaje o en la orientación académico-profesional, así como predictores con importancia específica diferencial en cada etapa. En el ámbito metodológico, el

---

Fecha de recepción del original: 28-03-2024.

Fecha de aprobación: 05-07-2024.

Cómo citar este artículo: Constante-Amores, A., Arroyo-Resino, D., Sánchez-Munilla, M., y Asensio-Muñoz, I. (2024). Contribución del *machine learning* al análisis de la repetición escolar en España: un estudio con datos PISA [Contribution of machine learning to the analysis of grade repetition in Spain: A study based on PISA data]. *Revista Española de Pedagogía*, 82 (289), 539-562. <https://doi.org/10.22550/2174-0909.4014>

estudio hace una aportación a la mejora de la especificación de los modelos predictivos.

**Palabras clave:** PISA, repetición escolar, aprendizaje automático, variables de contexto, regresión logística multinivel, educación obligatoria.

### Abstract:

**Introduction:** The rate of grade repetition is excessively high in Spain despite being a controversial measure. In order to obtain evidence to contribute to reducing it in compulsory education, the present work is an in-depth study of the PISA 2018 context indices that are most closely linked to this phenomenon. **Method:** With the sample of Spanish students ( $n = 35943$ ), we used an automatic machine learning method to select and order the predictors, and multilevel logistic regression (students and centres) to quantify the con-

tribution of each one. **Results:** For each educational stage we obtained the 30 most significant contextual variables, which explain 65.5% of the grade repetition variance in primary education and almost 55.7% in secondary education. **Conclusions:** The main indicators are principally at student level, which suggests the suitability of psychoeducational interventions based more on individual support than general policies. This gives rise to potentially more efficient and equitable measures than grade repetition, aimed at, for example, the management of learning time or academic/professional guidance, and predictors with specific differential significance at each stage. Methodologically, the study contributes to improving the specification of predictive models.

**Keywords:** PISA, grade repetition, machine learning, contextual variables, multilevel logistic regression, compulsory education.

## 1. Introducción

Como es sabido, la repetición de curso es una estrategia consistente en retener al alumnado en un curso de una determinada etapa educativa cuando este no demuestra haber alcanzado los niveles mínimos de conocimiento (Jimerson y Ferguson, 2007; López y García, 2020). Sin embargo, se trata de una medida controvertida, que algunos autores consideran positiva porque ayuda a los estudiantes con bajo rendimiento a mejorar sus competencias y a ganar madurez (Battisttin y Schizzerotto, 2019; Ikeda y García, 2014; Valbuena *et al.*, 2021), y otros, como una medida

meramente sancionadora (Schwerdt *et al.*, 2017). Además, aunque sobre el tema hay estudios de síntesis, revisiones sistemáticas y metaanálisis que abarcan prácticamente todo el siglo xx (Allen *et al.*, 2009; Jimerson 2001), existen dudas sobre su efectividad. Es más, parece ser, incluso, perjudicial al estar asociada con un descenso en el rendimiento de los estudiantes (Asensio *et al.*, 2018; Hattie, 2017; López *et al.*, 2018, Nieto-Isidro y Martínez-Abad, 2023) y vincularse con aspectos negativos como frustración (López-Rupérez *et al.*, 2021); sensaciones de discriminación (Van Canegem *et al.*, 2021); o disminución de la

autoestima, del autoconcepto académico o de la motivación, de forma tal que afecta a las expectativas académicas (González-Nuevo *et al.*, 2023; Mathys *et al.*, 2019; Peixoto *et al.*, 2016).

La normativa al respecto en España es sensible a esta realidad y, desde la Ley Orgánica 1/1990, del 3 de octubre, de Ordenación General del Sistema Educativo (LOGSE), se establece que la repetición es una medida de atención a la diversidad de carácter extraordinario. En la realidad, sin embargo, su uso no es tan excepcional. Según informa el Ministerio de Educación y Formación Profesional (2022), en la primera etapa de educación secundaria, su incidencia es del 8.7% y, en primaria, el 2.4% del alumnado es repetidor, porcentaje solo superado en la Unión Europea por Portugal (3.6%), República Eslovaca (2.9%) y Austria (2.9%). Por otra parte, con datos del Programa para la Evaluación Internacional de Alumnos (PISA) de 2009, 2015 y 2018, López y García (2020), que comparan el porcentaje de alumnos de 15 años que han repetido, al menos, una vez en la Unión Europea, concluyen que España, aunque ha reducido su tasa de 2009 a 2018 de forma notable, en este último año alcanza todavía el 28.71%. Esto sitúa al país en el tercer lugar, solo por detrás de Luxemburgo (32.02%) y Bélgica (29.06%), cuando la media de la Organización para la Cooperación y el Desarrollo Económico (OCDE) está en 11.4%, y la de Finlandia, por ejemplo, en 3.2%.

A la vista de lo expuesto, resulta necesario seguir investigando este fenómeno

con el fin último de aportar información para poder actuar sobre él con base en la evidencia, no solo desde la política, sino también desde la práctica educativa. Así, se espera contribuir a optimizar el dato, de manera que España pueda llegar a aproximarse, por lo menos, a la media de la OCDE.

En este sentido, el presente estudio se sitúa en una línea de búsqueda de predictores a través de la explotación secundaria de los datos de PISA, de los que ya existe investigación derivada de ediciones anteriores que muestra que la repetición de curso es multicausal (Arroyo *et al.*, 2019; Carabaña, 2013; Cordero *et al.*, 2014).

Con base en la clasificación de los cuestionarios de contexto que utiliza PISA (OCDE, 2019), representada en la Figura 1, a continuación, se expone un breve resumen del estado de la cuestión en torno a las variables más relevantes de los tres constructos.

En lo referido a los antecedentes, hay evidencia, sobre todo, de que el alumnado repetidor suele pertenecer a contextos socioeconómicos desfavorecidos (López y García, 2020). Sin embargo, la condición de inmigrante muestra resultados no concluyentes, ya que Cordero *et al.* (2014), García-Pérez *et al.* (2014) o Warren *et al.* (2014) indican que los inmigrantes de primera generación (nacidos en el extranjero) tienen una mayor probabilidad de repetir que los nativos, mientras que, por ejemplo, en el estudio de Choi *et al.* (2018), esta variable no resulta significativa.

FIGURA 1. Estructura modular de las variables de contexto en PISA 2018.

Competencia lectora	Constructos de antecedentes de los estudiantes	Constructos asociados a la escuela			Constructos no cognitivos y metacognitivos
<p>6. Estatus socioeconómico del estudiante y la familia</p> <p>7. Migración y cultura</p>	<p>5. Experiencia científica fuera de la escuela</p> <p>8. Recorrido educativo en la primera infancia</p>	ENSEÑANZAS Y APRENDIZAJE			<p>4. Medidas relacionadas con la lectura: actitudes, motivación y estrategias</p>
		1. Cualificación y conocimiento profesional del profesor	2. Práctica de enseñanza de las ciencias	11. Tiempo de aprendizaje y currículo	
Categorías generales		POLÍTICAS ESCOLARES			<p>10. Disposiciones para la competencia global</p>
		13. Implicación de los padres	3. Entorno de aprendizaje escolar de la lectura	14. Contextos y recursos escolares	
		12. Clima escolar: relaciones interpersonales, confianza y expectativas	GOBERNANZA	16. Evaluación del estudiante, evaluación institucional y responsabilidad	

Fuente: OCDE (2019).

En segundo lugar, entre los constructos asociados a la escuela, se han estudiado, fundamentalmente, el tiempo de aprendizaje, el clima escolar o el currículo, que se asocian con la repetición (OCDE, 2019). Así, Asensio *et al.* (2018) y López *et al.* (2018) demuestran la importancia del número de clases a la semana o de la duración de las clases en la repetición, mientras que Seabra y Ferrão (2016) hallan una relación significativa entre la repetición de curso y los problemas de disciplina en los centros educativos, con un impacto del clima disciplinario más fuerte que el del nivel socioeconómico. Por otro lado, Arroyo *et al.* (2019) vinculan la repetición con variables curriculares, como haber cursado ciencias en el año anterior o haber elegido alguna asignatura optativa de esta rama.

En el grupo de constructos no cognitivos y metacognitivos (en adelante CN-CyM), la motivación, la autoestima, las expectativas académicas futuras y el sentido de pertenencia al grupo son variables de gran relevancia (Fernández-Lasarte *et al.*, 2019; Hornstra *et al.*, 2017; Van Canegem *et al.*, 2021), de manera que un desarrollo positivo de estos aspectos minimiza el efecto de la repetición (Marsh *et al.*, 2018). Igualmente, el autoconcepto académico y la orientación hacia metas tienen una elevada capacidad predictiva del fenómeno que nos ocupa (Ferla *et al.*, 2009; Rodríguez-Rodríguez, 2022).

A pesar del número de investigaciones existentes, una limitación que dificulta perfilar un estado de la cuestión más riguroso se deriva del hecho de que los

predictores se suelen introducir en los modelos de manera arbitraria, quizá por la inexistencia de referentes teóricos que permitan, en materia educativa, una modelización que procure minimizar errores de especificación. Y ello aunque el principal supuesto de los modelos multinivel (que son los que más se emplean en la explotación secundaria de los datos de PISA por su isomorfismo con la realidad que se pretende modelizar) sea que estén bien especificados, esto es, que cumplan con la condición de no dejar «fuera del modelo predictores importantes» (Gaviria y Castro, 2005, p. 86).

Ante ello, cobra interés explorar modelos predictivos en los que la disminución de los errores estadísticos venga acompañada del intento de reducir al mínimo también los errores de especificación. En esta línea se sitúa la presente investigación junto con otras que emplean la minería de datos; se basan en el aprendizaje automático, o *machine learning* (Urbina y De la Calleja, 2017), tanto a nivel internacional (Gamazo y Martínez-Abad, 2020; Kılıç *et al.*, 2017; Kiray *et al.*, 2015; Liu y Ruiz, 2008; Martínez-Abad y Chaparro, 2017; Martínez-Abad *et al.*, 2020) como centradas en muestras españolas (Arroyo *et al.*, 2024a; Arroyo *et al.*, 2024b; Asensio *et al.*, 2018).

Otra limitación está en que la investigación anterior no suele diferenciar entre estudiantes de las dos etapas obligatorias, centrándose la mayoría de los estudios en el alumnado de secundaria, lo que dificulta que emerjan variables con importancia diferencial en educación primaria.

Con el fin de contribuir a la superación de las limitaciones mencionadas, el objetivo general que se persigue aquí es obtener información relevante para una modelización más ajustada de la repetición de curso en España, en las dos etapas obligatorias, a partir de las bases de datos de PISA 2018 (Ministerio de Educación y Formación Profesional, 2019), que aportan una ingente información sobre variables de contexto (OCDE, 2019).

Como objetivos específicos, se pretende:

- Identificar el algoritmo de *machine learning* [árboles de decisión individuales (CART y C5.0), *random forest* y *stochastic gradient boosting* (GBM)] más preciso y, con él, seleccionar aquellas variables más vinculadas con la repetición de curso en las dos etapas y ordenarlas según su importancia diferencial en cada una.
- Determinar la contribución relativa de cada factor (constructos de antecedentes del estudiante asociados a la escuela y CNCyM) en la probabilidad de que un estudiante repita curso en primaria y secundaria teniendo en cuenta la estructura en dos niveles de los datos (estudiantes y centros).

## 2. Método

Se realiza un análisis secundario de los datos de la evaluación internacional de PISA 2018 mediante un diseño no experimental, cuantitativo y transversal de tipo correlacional, de acuerdo con el manual de recomendaciones técnicas de PISA (OCDE, 2019).

### 2.1. Participantes

Con muestreo bietápico, la muestra del alumnado español que participa en PISA 2018 es de 35 943, lo que supone una cobertura del 92% de la población nacional de 15 años (Ministerio de Educación y Formación Profesional, 2019), con 49,96% de mujeres y una edad media de 15.84 años ( $DT = 0.29$ ). Según el tipo de centro, el 65% son estudiantes matriculados en centros públicos; el 28%, en concertados, y el 7%, en privados. La información completa se puede consultar en el segundo anexo del documento de la OCDE (2020).

### 2.2. Instrumentos

Todas las variables se miden a través de los cuestionarios de contexto de PISA 2018 para estudiantes y para directores (OCDE, 2019). Las variables dependientes son la repetición en primaria y en secundaria, con una incidencia de casi el doble en esta última etapa según informan los propios participantes (Tabla 1).

En cuanto a las variables independientes, se introducen, en la fase inicial, 85 (consultar Tablas S1 y S2 en <https://acortar.link/hfDnvO>). De este conjunto de predictores, 83 son índices compuestos (WLE y SUM), de los cuales 67 índices fueron construidos a partir de las respuestas de los estudiantes (nivel 1) y 17 a partir de las respuestas de los directores de centros educativos (nivel 2). Las dos variables restantes son ítems simples que se han incluido en el modelo por la abundante evidencia previa existente: *género del estudiante* (nivel 1) y *porcentaje de estudiantes de 15 años procedentes de hogares socioeconómicamente desfavorecidos* (nivel 2).

TABLA 1. Repetición de curso en la educación obligatoria en España.

<b>¿Has repetido alguna vez en la educación primaria?</b>				
	<i>n</i>	%	<i>n</i> (ponderado)	% (ponderado)
No	30 078	89.927	340 287	88.606
Sí, una vez o más	3 369	10.073	43 756	11.394
Total válidas	33 447			
<b>¿Has repetido alguna vez en la educación secundaria?</b>				
	<i>n</i>	%	<i>n</i> (ponderado)	% (ponderado)
No	27 902	80.354	310 497	77.313
Sí, una vez o más	6 822	19.646	91 114	22.687
Total válidas	34 724			

Fuente: elaboración propia a partir de los datos de España en PISA (2018).

Posteriormente, en los modelos multinivel se incluyen solo los que, a través del algoritmo de *machine learning* más preciso, muestren mayor vinculación con la repetición en cada una de las etapas estudiadas.

### 2.3. Procedimiento

Un resumen detallado del procedimiento de aplicación de los cuestionarios puede verse en el informe español (Ministerio de Educación y Formación Profesional, 2019).

### 2.4. Análisis de datos

Se diferencian dos fases, de modo que primero se realiza un estudio descriptivo y exploratorio (Rosenthal y Rosnow, 2008), mediante algoritmos supervisados de *machine learning*, y, una vez seleccionadas y ordenadas las variables más importantes, se prueban varios modelos predictivos de acuerdo con la estructura jerárquica de los datos (nivel 1: estudiante y nivel 2: centro escolar).

Previo al algoritmo de *machine learning*, se preprocesan los datos, como recomienda Raschka (2015), se tratan los valores perdidos mediante la técnica de imputación múltiple con ecuaciones encadenadas (Rivero, 2011) y se procede a la división aleatoria de los casos en dos conjuntos: uno de entrenamiento (70%), con el que se ajusta el modelo, y otro de validación (30%), con el que se evalúa el modelo a través del área bajo la curva (AUC) por tratarse de dos variables dependientes con niveles desbalanceados. También se realiza un centrado y escalado de los predictores continuos a fin de evitar la sobrerrepresentación de las variables con magnitud mayor y no se obtiene ningún predictor con varianza cero o próxima a cero.

Para responder al primer objetivo, se compara el rendimiento de cuatro algoritmos [árboles de decisión individuales (CART y C5.0), *random forest* y *stochastic gradient boosting*] a través del AUC con



el método de *grid-search*, basado en la validación cruzada con 10 pliegues y en la optimización de los hiperparámetros más importantes de manera conjunta. Con el método de ordenación por importancia relativa, *stochastic gradient boosting* resulta ser el algoritmo más preciso si se atiende al test de Friedman aplicado en la muestra de entrenamiento. En aras de la parsimonia, dado que una limitación de dicho algoritmo es que no sugiere un punto de corte fijo en cuanto a la selección de un número óptimo de predictores (Sarkar *et al.*, 2018), se decide, siguiendo a Gorostiaga y Rojo-Álvarez (2016), examinar los siguientes conjuntos de variables: 15, 20, 25, 30 y 35. Así, se aporta una mayor precisión en la selección del conjunto óptimo, que resulta ser, tanto en primaria como secundaria, el formado por 30 de todos los índices incluidos inicialmente.

Para lograr el segundo objetivo, se utiliza la técnica de regresión logística binaria multinivel, ya que la varianza aleatoria del modelo nulo resulta estadísticamente significativa y el coeficiente de correlación intraclase (CCI) es superior al 10% (Lee, 2000). Con base en el supuesto de multicolinealidad, se eliminan *posesiones del hogar* (HOMEPOS), *educación de la madre* (MISCED) e *índice de la situación profesional más alta de los padres* (HISEI) por estar implícitos en el *índice de estatus económico, social y cultural* (ESCS); y el predictor *riqueza familiar* por su alta correlación con este índice.

A continuación, como exige el análisis multinivel, se suprimen los centros con menos de 20 sujetos, con lo que se exclu-

yen 1532 alumnos pertenecientes a 133 centros, por lo que la muestra final es de 34411 estudiantes.

Se generan un total de cinco modelos: el modelo nulo, sin predictores; el modelo 1, formado por las variables de antecedentes del estudiante; el modelo 2, compuesto por las variables asociadas a la escuela; el modelo 3, con variables de CNCyM; y el modelo 4, con todos los predictores. Así, se puede analizar la contribución de los factores de escuela y CNCyM una vez controlados los antecedentes del estudiante. Las variables se introducen en los modelos según el orden obtenido con el algoritmo de *machine learning* (Arroyo *et al.*, 2019; Constante *et al.*, 2021). Finalmente, para evaluar el ajuste, se utilizan los índices AIC, BIC y Deviance, un estadístico de reducción que permite comparar modelos anidados estimando el porcentaje de varianza reducido ( $R^2$ ), y se calcula su significatividad (Cameron y Windmeijer, 1997). La interpretación de los resultados se basa en el porcentaje de varianza explicada (PVE), que cuantifica la capacidad explicativa del modelo final, y en las *odds ratios*, en las que valores superiores a 1 indican que la variable aumenta la probabilidad de repetir curso (actúa como un factor de riesgo), mientras que valores inferiores a 1 sugieren que disminuye esta probabilidad (funcionan como un factor protector).

Como *software*, se utilizan los paquetes de R versión 4.4.0. *caret*, *stochastic gradient boosting* y *lme4* para comparar los algoritmos de *machine learning*, la selección de las variables más importantes y la regresión logística binaria multinivel, respectivamente.

### 3. Resultados

De los algoritmos sometidos a prueba, *stochastic gradient boosting* y *random forest* resultan ser los más precisos tanto en primaria (Figura 2) como en secundaria

(Figura 3). Con el test de Friedman, destaca *stochastic gradient boosting*, que es el que finalmente se usa para seleccionar y jerarquizar las variables más importantes asociadas a la repetición en educación obligatoria.

FIGURA 2. Comparación de algoritmos de *machine learning* para educación primaria.

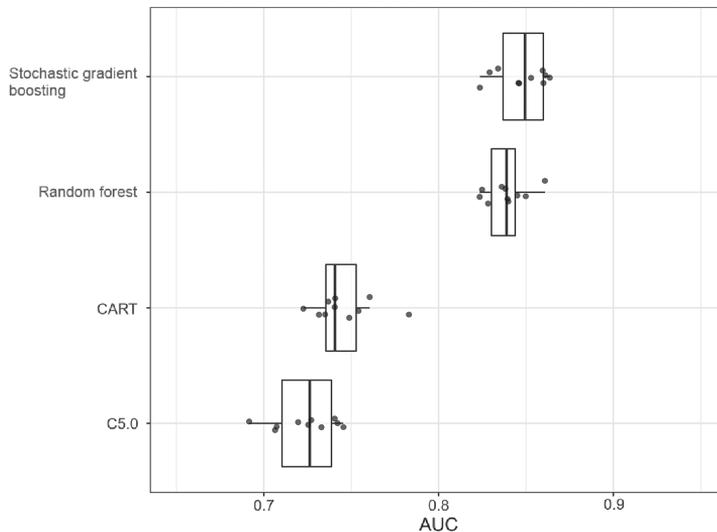
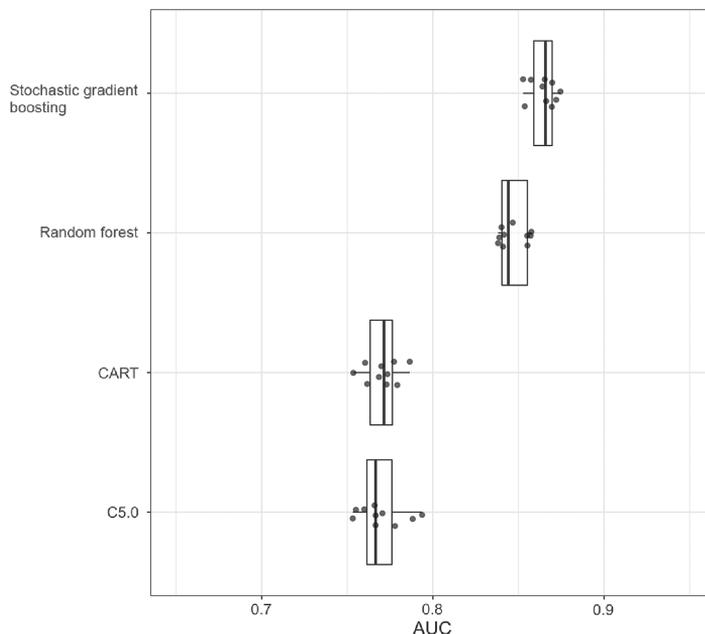


FIGURA 3. Comparación de algoritmos de *machine learning* para educación secundaria.



En la Tabla 2 se ofrece el resultado empírico que aporta *stochastic gradient boosting*, en el que, de las 30 variables más importantes en primaria, 12 son de antecedentes, 9 asociadas a la escuela y 9 de CN-CyM. Todas se corresponden con el nivel 1, excepto el *porcentaje de alumnos procedentes de hogares socioeconómicamente*

*desfavorecidos*, que es de nivel 2 y ocupa el puesto 23.º. En esta etapa educativa, los predictores más relevantes son los siguientes: primero, el *índice de estatus económico, social y cultural (ESCS)*; segundo, la *situación profesional esperada del alumno (BSMJ)*; y tercero, el *tiempo dedicado al aprendizaje de matemáticas (MMINS)*.

TABLA 2. Variables más importantes asociadas a la repetición de curso en la educación primaria.

	<b>Etiqueta</b>	<b>Orden</b>	<b>Variables</b>
<b>Antecedentes del estudiante</b>	ESCS	1.º	Índice de estatus económico, social y cultural
	HOMEPOS	4.º	Posesiones del hogar
	SCCHANGE	5.º	Número de cambios de centro
	HISEI	8.º	Índice de la situación profesional más alta de los padres
	MISCED	9.º	Educación de la madre
	IMMIG	12.º	Condición de inmigrante
	ICTRES	16.º	Recursos TIC del hogar
	HEDRES	17.º	Recursos educativos en el hogar
	DURECEC	20.º	Duración de la educación y atención en la primera infancia
	SC048Q03NA	23.º	Porcentaje de alumnos procedentes de hogares socioeconómicamente desfavorecidos
	BMMJ1	27.º	ISEI de la madre
	WEALTH	30.º	Riqueza familiar
<b>Asociados a la escuela</b>	MMINS	3.º	Tiempo de aprendizaje de matemáticas (minutos por semana)
	LMINS	6.º	Tiempos de aprendizaje de lengua (minutos por semana)
	INFOCAR	10.º	Información sobre las carreras
	SMINS	11.º	Tiempo de aprendizaje en ciencias (minutos por semana)
	DIRINS	13.º	Enseñanza dirigida por el profesor
	TMINS	14.º	Tiempo de aprendizaje en ciencias (minutos por semana)
	PERFEED	18.º	<i>Feedback</i> percibido por parte del profesorado
	DISCLIMA	28.º	Clima disciplinario en las clases de lengua (centrales en esta edición de PISA)
BEINGBULLIED	29.º	Experiencia de acoso escolar	

No cognitivos y metacognitivos	BSMJ	2.º	Situación profesional esperada del alumno
	PISADIFF	7.º	Percepción de dificultad de la prueba PISA
	METASUM	15.º	Metacognición resumen
	STUBMI	19.º	Índice de masa corporal del alumno
	GCESELFEEFF	21.º	Autoeficacia en cuestiones globales
	BELONG	22.º	Bienestar subjetivo: sentimiento de pertenencia al centro educativo
	SCREADDIFF	24.º	Autoconcepto lector: percepción de la dificultad
	GCAWARE	25.º	Conciencia del alumno sobre cuestiones globales
COGFLEX	26.º	Flexibilidad/adaptabilidad cognitiva del estudiante	

En la Tabla 3 se ofrecen las variables que emergen como más importantes con *stochastic gradient boosting* en secundaria: 8 de antecedentes del estudiante, 14 asociados a la escuela y 8 no cognitivos y metacognitivos del estudiante. En esta etapa, también hay 29 de nivel 1 y solo una de nivel 2: *comportamiento de los alumnos que dificulta el aprendizaje* (puesto 11.º). Los predictores más re-

levantes aquí resultan ser la *situación profesional esperada del alumno* (BSMJ), el *tiempo de aprendizaje dedicado a las ciencias* (SMINS) y el *índice de estatus económico, social y cultural* (ESCS). Los análisis descriptivos de las 30 variables más importantes presentes en la Tabla 2 y en la Tabla 3 se encuentran en el siguiente enlace: <https://acortar.link/hfDnvO> (Tablas S3 y S4).

Tabla 3. Variables más importantes asociadas a la repetición de curso en la educación secundaria.

Área	Etiqueta	Orden	Variables
Antecedentes del estudiante	ESCS	3.º	Índice de estatus económico, social y cultural
	CHANGE	5.º	Número de cambios en la biografía educativa
	HOMEPOS	9.º	Posesiones del hogar
	BMMJ1	10.º	ISEI de la madre
	HISEI	12.º	Índice de la situación profesional más alta de los padres
	HEDRES	20.º	Recursos educativos del hogar
	ICTRES	21.º	Recursos TIC del hogar
	DURECEC	23.º	Duración de la educación y atención en la primera infancia



Asociados a la escuela	SMINS	2.º	Tiempo de aprendizaje en ciencias (minutos por semana)
	LMINS	4.º	Tiempo de aprendizaje de lengua (minutos por semana)
	MMINS	6.º	Tiempo de aprendizaje en matemáticas (minutos por semana)
	INFOCAR	7.º	Información sobre las carreras
	STUBEHA	11.º	Comportamiento de los alumnos que dificulta el aprendizaje
	HOMESCH	13.º	Uso de las TIC en casa para actividades de trabajo escolar
	TMINS	15.º	Tiempo de aprendizaje en total (minutos por semana)
	ICTOUTSIDE	17.º	Uso de las TIC relacionado con las asignaturas fuera de clase
	ENTUSE	18.º	Uso de las TIC relacionado con el ocio fuera de clase
	INFOJOB2	19.º	Información sobre el mercado laboral facilitada fuera de la escuela
	PERFEED	22.º	<i>Feedback</i> percibido por parte del profesorado
	INFOJOB1	24.º	Información sobre el mercado laboral proporcionada por el centro escolar
	DISCRIM	27.º	Discriminación del clima escolar
DIRINS	29.º	Enseñanza dirigida por el profesor	
No cognitivos y metacognitivos	BSMJ	1.º	Situación profesional esperada del alumno
	MASTGOAL	8.º	Orientación hacia objetivos de dominio
	PISADIFF	14.º	Percepción de la dificultad de la prueba PISA
	GCELFEFF	16.º	Autoeficacia en cuestiones globales
	RESILIENCE	25.º	Resiliencia del alumnado
	METASUM	26.º	Metacognición: resumen
	STUBMI	28.º	Índice de masa corporal del alumno
GCAWARE	30.º	Conciencia de los alumnos sobre cuestiones globales	

Para el logro del segundo objetivo, una vez identificadas y ordenadas las variables que tienen una mayor relación con la repetición, se procede a especificar el modelo predictivo, para lo cual se dicotomizan las

variables independientes de naturaleza categórica según indican Pardo y Ruiz (2013) (Tabla 4). Con el predictor DURECEC, se siguen las recomendaciones de la OCDE (2019).

TABLA 4. Recodificación de variables independientes para el modelo de regresión logística multinivel.

Etiqueta	Variable independiente categórica	Valores recodificados
SCCHANGE	Número de cambios de colegio	0 = Ningún cambio 1 = Uno o más cambios
CHANGE	Número de cambios en la biografía educativa	
IMMIG	Condición de inmigrante	0 = Inmigrante 1 = Nativo
DURECEC	Duración de la educación y atención en la primera infancia	0 = Tres o menos 1 = Cuatro o más

En la Tabla 5 se resumen los modelos de regresión de la repetición en primaria; se identifica con N1 o N2 tras el nombre de la variable el nivel al que pertenece cada una. En el modelo 1, todos los antecedentes del estudiante resultan estadísticamente significativos y explican el 62.380% de la variabilidad de la repetición. Atendiendo a las *odds ratio*, el predictor *número de cambios del colegio* (SCCHANGE) es el que supone mayor probabilidad de repetición. Hay que destacar que *el porcentaje de alumnos procedentes de hogares socioeconómicamente desfavorecidos* (SC048Q03NA) es la única variable significativa del nivel 2, y el *ISEI de la madre* (BMMJ1) tiene unas *odds ratio* muy próximas a 1. El modelo 2, constituido por las variables asociadas a la escuela, explica el 19% de la varianza del criterio y, aunque todas son significativas, cuatro tienen unas *odds ratio* de 1. De ellas, la *experiencia de acoso escolar* (BEINGBULLIED) y la *enseñanza dirigida por el profesor* (DIRINS) son los principales factores de riesgo. El modelo 3 explica en torno al 25% de la varianza. Incluye únicamente CNCyM, de los cuales la *conciencia que tiene el alumno*

*sobre cuestiones globales* (GCAWARE) no resulta significativa y tiene una *odd ratio* muy próxima al *índice de masa corporal* (STUBMI); en cambio, la variable con mayor efecto es el *autoconcepto lector: dificultad percibida* (PISADIFF). El último modelo (M4) explica el 61% de la varianza de la repetición de curso en educación primaria. En él, ni los efectos del *tiempo de aprendizaje en ciencias (minutos por semana)* (SMINS) ni el *índice de masa corporal* (STUBMI) son significativos. Además, hay ocho variables con *odds ratio* muy próximas a 1; de ellas, las de mayor efecto son *número de cambios de colegio* (SCCHANGE), *enseñanza dirigida por el profesor* (DIRINS) y *feedback percibido por el profesorado* (PERFEED).

Por último, como se observa en la Tabla 5, en lo que concierne al ajuste, el modelo 4 es el que tiene un menor AIC y BIC, con una reducción importante de la varianza equivalente a un  $R^2$  del 9% respecto al modelo 1. También es significativa respecto del modelo nulo, con una reducción de la varianza del modelo 1 (8%), del modelo 2 (5%) y del modelo 3 (9%).

TABLA 5. *Odds ratio* y errores típicos para los modelos de regresión logística multinivel en primaria.

	M0 (Nulo)	M1 (Antecedentes)	M2 (Escuela)	M3 (CNCyM)	M4 (Todas)
Efectos fijos					
Intercepto	0.10(0.03)***	0.14(0.08)***	0.07(0.09)***	0.41(0.11)***	0.38(0.17)***
ESCS_N1		0.59(0.03)***			0.67(0.03)***
SCCHANGE_N1		2.10(0.04)***			1.89(0.04)***
IMMIG_N1		0.62(0.05)***			0.58(0.05)***
ICTRES_N1		0.90(0.03)***			0.90(0.03)***
HEDRES_N1		0.89(0.02)***			0.93(0.03)***
DURECEC_N1		0.69(0.05)***			0.77(0.05)***
SC048Q03NA_N2		1.01(0.00)***			1.01(0.00)***
BMMJ1_N1		0.99(0.00)***			1.00(0.00)***
MMINS_N1			1.00(0.00)***		1.00(0.00)***
LMINS_N1			1.00(0.00)***		1.00(0.00)***
INFOCAR_N1			0.72(0.02)***		0.79(0.02)***
SMINS_N1			1.00(0.00)***		1.00(0.00)***
DIRINS_N1			1.30(0.03)***		1.24(0.02)***
TMINS_N1			1.00(0.00)***		1.00(0.00)***
PERFEEED_N1			1.25(0.03)***		1.22(0.02)***
DISCLIMA_N1			0.83(0.02)***		0.89(0.02)***

BEINGBULLIED_N1	1.30(0.02)***			1.13(0.02)***
BSMJ_N1		0.97(0.00)***		0.98(0.00)***
PISADIFF_N1		1.31(0.02)***		1.19(0.02)***
METASUM_N1		0.69(0.02)***		0.75(0.02)***
STUBMI_N1		1.01(0.00)***		1.00(0.00)***
GCSELEFF_N1		0.87(0.02)***		0.89(0.02)***
BELONG_N1		0.87(0.02)***		0.93(0.02)***
SCREADDIFF_N1		1.89(0.02)***		1.08(0.02)***
GCAWARE_N1		0.99(0.02)		1.01(0.03)***
COGFLEX_N1		1.08(0.02)***		1.06(0.02)***
Efectos aleatorios				
Varianza	0.52(0.05)***	0.20(0.03)***	0.42(0.04)***	0.39(0.04)***
PVE (%)	62.38	18.62	25.53	0.61
Índices de ajuste				
AIC	22198.69	19804.01	21053.05	18072.87
BIC	22215.58	19888.48	21145.96	18309.368
Loglik	-11097.34	-9892.01	-10515.53	-9008.43
Deviance	20724.84	19026.15	19749.16	17292.51
Chisq		2410.72	1163.65	1767.1
Pr (> Chisq)		0.00	0.00	0.00

Nota: \* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$ . La categoría de referencia para interpretar las *odds ratio* es, en todos los casos, el valor 1.

Finalmente, en la Tabla 6 se ofrecen los modelos de regresión de la repetición en secundaria; se identifica con N1 o N2 tras el nombre de la variable el nivel al que pertenece cada una. El modelo 1 explica, aproximadamente, el 51% del criterio. En él, el efecto de *recursos TIC del hogar* (ICTRES) resulta no significativo, mientras que el principal factor de riesgo es el *número de cambios en la biografía educativa* (CHANGE). El modelo 2 explica el 32% de la varianza; todas las variables son estadísticamente significativas, aunque cuatro tienen unas *odds ratio* muy próximas a 1. El mayor efecto corresponde al *comportamiento de los alumnos que dificulta el aprendizaje* (STUBEHA), variable de nivel 2. El modelo 3, con predictores CNCyM, explica el 14%. La *resiliencia del alumna-do* (RESILIENCE) no resulta significativa y hay otros dos predictores con *odds ratio* muy próximas a 1. La variable con mayor factor de riesgo es el *autoconcepto lector*:

*dificultad percibida* (PISADIFF). En el modelo global, el *tiempo de aprendizaje en ciencias* (SMINS), el *tiempo de aprendizaje en matemáticas* (MMINS) y la *resiliencia* (RESILIENCE) no resultan significativas y hay otros seis predictores con *odds ratio* muy próximos a 1. Este conjunto de predictores explica casi el 56% de la repetición de curso en secundaria. Las variables con mayor efecto son *número de cambios del colegio* (SCCHANGE), *comportamiento de los alumnos que dificulta el aprendizaje* (STUBEHA) y *enseñanza dirigida por el profesor* (DIRINS).

Por último, cabe destacar que el modelo 4 es el que mejor ajusta, con una reducción significativa de la varianza equivalente a un R<sup>2</sup> del 10.298% respecto al modelo 1, aunque también es significativa la reducción del modelo 1 respecto al modelo 0 (7%) y la de los modelos 2 y 3 respecto al nulo (5% y 10%, respectivamente).

TABLA 6. *Odds ratio* y errores típicos para los modelos de regresión logística multinivel en secundaria.

	M0 (Nulo)	M1 (Antecedentes)	M2 (Escuela)	M3 (CNCyM)	M4 (Todas)
Efectos fijos					
Intercepto	0.21(0.03)***	0.27(0.06)***	0.14(0.07)***	1.00(0.10)	0.70(0.13)***
ESCS_N1		0.67(0.02)***			0.76(0.02)***
CHANGE_N1		2.44(0.03)***			2.24(0.03)***
BMMJ1_N1		0.98(0.01)***			0.99(0.00)***
HEDRES_N1		0.85(0.02)***			0.92(0.02)***
ICTRES_N1		0.91(0.02)			0.91(0.02)***
DURECEC_N1		0.69(0.05)***			0.76(0.04)***
SMINS_N1			1.00(0.00)***		1.00(0.00)***
LMINS_N1			1.01(0.00)***		1.00(0.00)***

## Contribución del *machine learning* al análisis de la repetición escolar en España: un estudio con...

MMINS_N1			1.00(0.00)***		1.00(0.00)
INFOCAR_N1			0.72(0.02)***		0.76(0.02)***
STUBEHA_N2			1.42(0.03)***		1.25(0.02)***
HOMESCH_N1			0.91(0.02)***		0.93(0.02)***
TMINs_N1			1.00(0.00)***		1(0.00)***
ICTOUTSIDE_N1			0.89(0.02)***		0.96(0.02)***
ENTUSE_N1			1.10(0.02)***		1.13(0.02)***
INFOJOB2_N1			1.13(0.02)***		1.11(0.02)***
PERFEED_N1			1.16(0.02)***		1.15(0.02)***
INFOJOB1_N1			0.84(0.02)***		0.85(0.02)***
DISCRIM_N1			1.39(0.02)***		1.20(0.02)***
DIRINS_N1			1.21(0.02)***		1.21(0.02)***
BSMJ_N1				0.97(0.00)***	0.98(0.00)***
MASTGOAL_N1				0.78(0.02)***	0.82(0.02)***
PISADIFF_N1				1.23(0.02)***	1.14(0.02)***
GCELFEEFF_N1				0.86(0.02)***	0.88(0.02)***
RESILIENCE_N1				1.00(0.02)	0.99(0.02)
METASUM_N1				0.72(0.02)***	0.79(0.02)***
STUBMI_N1				1.01(0.00)***	1.01(0.00)***
GCAWARE_N1				1.08(0.02)***	0.76(0.02)***
Efectos aleatorios					
Varianza	0.70(0.04)***	0.34(0.05)***	0.47(0.04)***	0.60(0.05)***	0.32(0.04)***
PVE (%)		51	32.43	14	55.69
Índices de ajuste					
AIC	32841,56	29736,77	30809,14	29611,70	26708,20
BIC	32858.46	29804.34	30944.28	29696.16	26961.58
Loglik	-16418.78	-14860.38	-15388.57	-14795.85	-13324.10
Deviance	30681.45	28238.71	29013.64	27662.13	25330.76
Chisq		3116.8	2060.3	152.57	6215.6
Pr (> Chisq)	0.70(0.04)***	0.34(0.05)***	0.47(0.04)***	0.60(0.05)***	0.32(0.04)***

Nota: \* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$ . La categoría de referencia para interpretar las *odds ratio* es, en todos los casos, el valor 1.

#### 4. Discusión

En la modelización a la que se ha llegado, los índices de contexto más relevantes resultan ser, en su mayoría, del primer nivel; esto es, se imponen las características de los estudiantes sobre las de los centros. Este resultado coincide con el obtenido por López *et al.* (2023) en su revisión sistemática de investigaciones realizadas con estudiantes españoles. También con los de otros análisis secundarios en los que el objeto de estudio es el rendimiento, como el que realizan Choi *et al.* (2018) con datos de estudiantes españoles en PISA 2012 y en PIRLS 2006 (Progress in International Reading Literacy Study), o Lopes *et al.* (2022) con datos de estudiantes portugueses en PIRLS 2016.

Cabe concluir que las tres variables más vinculadas con la repetición en toda la educación obligatoria son el *estatus social*, del grupo de antecedentes del estudiante; el *tiempo de aprendizaje*, del grupo de constructos asociados a la escuela; y el *nivel de aspiraciones*, de los CNCyM, variando el orden entre las dos etapas. En esta última, los resultados van en la línea de lo encontrado por Asensio *et al.* (2018) y López *et al.* (2018), que, sin diferenciar por etapa, sitúan las expectativas entre las variables más importantes respecto al rendimiento. En cuanto a la importancia de la variable de antecedentes del estudiante, el presente estudio está en línea con lo obtenido por Carabaña (2013), Choi *et al.* (2018) o García-Pérez *et al.* (2014). Sin embargo, aquí se aporta información específica relativa al peso diferencial según la etapa, que puede interpretarse como un indicador de que el sistema educativo cumple con su función

compensatoria al pasar el ESCS a ocupar el tercer puesto en secundaria, detrás de una variable de orden psicológico y otra curricular. Esta evidencia es, sin duda, una aportación interesante de este estudio. En las dos etapas emergen otros elementos de carácter psicoeducativo relevantes como son, del grupo de variables asociadas a la escuela, la *información académico-profesional* (décima posición en primaria y séptima en secundaria), la *enseñanza dirigida* (décimo tercera en primaria y vigésima novena en secundaria), el *feedback percibido por parte del profesorado* y el *clima de clase*; y, del grupo CNCyM, la *dificultad de la tarea percibida*.

De manera específica, solo en primaria destaca la relación de la repetición con el *acoso* y con la *flexibilidad o adaptabilidad cognitiva*; en secundaria, con el *uso de las TIC*, con la *orientación a metas de dominio* y con la *conciencia sobre cuestiones globales*. Así, se incrementa en esta etapa el número de las variables asociadas a la escuela que llegan a ser importantes y disminuyen las de antecedentes, sobre las que el centro escolar tiene menos poder de actuación directa, lo que confirma y amplía los hallazgos de Arroyo *et al.* (2019) con datos de PISA 2015.

La emergencia de modelos diferenciales en las dos etapas permite concluir que el tratamiento de la repetición a nivel práctico y político debe tener en cuenta este aspecto. También la existencia de vínculos claros entre variables de antecedentes y repetición tiene implicaciones para la práctica al advertir de la necesidad de perseverar en prácticas compensatorias que mejoren la equidad, en especial

en primaria. Por otra parte, la importancia del tiempo dedicado al aprendizaje apunta a prácticas dirigidas a una gestión eficiente de esta variable, sin limitarse a incrementar el tiempo de estudio en las clases desfavorecidas, que es el problema que plantean, por ejemplo, las políticas de recuperación estival de las «deudas educativas» (Battistin y Schizzerotto, 2019). La solución más equitativa debería apoyarse en los principios de individualización, prematuridad (Choi *et al.*, 2018) e inclusión (Tapia y Álvarez, 2022), donde la formación del profesorado y la enseñanza colaborativa o la codocencia se presentan como estrategias prometedoras en los sistemas educativos más exitosos (Niemi, 2015). En cuanto al nivel de aspiraciones, resulta relevante trabajar la motivación tanto en primaria como en secundaria (Rhodes *et al.*, 2018; Rodríguez-Rodríguez, 2022). Otros aspectos específicos sobre los que habría que trabajar para disminuir la tasa de repetición en las dos etapas son la retroalimentación del aprendizaje, el clima o la orientación académico-profesional.

La disponibilidad de una base de datos tan amplia como PISA ha permitido, en este trabajo, el uso de algoritmos de aprendizaje automático. Ello constituye otra de sus principales aportaciones, ya que el estudio multinivel realizado se apoya en una selección previa que contribuye a minimizar los errores de especificación tan difícilmente evitables en investigación educativa. El empleo de técnicas de *machine learning* complementa y aumenta la sensibilidad de los modelos de regresión jerárquica y aporta criterios empíricos para quedarse con los predictores que importan y descartar los

irrelevantes. Aunque en esta segunda función no ha demostrado ser tan eficaz por la existencia de *odds ratios* significativas, pero iguales o muy próximas a 1, se trata de una metodología que permite una estimación cuantitativa no solo de la valía de cada predictor, sino, fundamentalmente, de la validez predictiva del modelo.

Como limitaciones, cabe advertir que los métodos de aprendizaje automático pueden sobreajustar los modelos a las características de la muestra actual. Además, los cuestionarios de contexto PISA se basan en autoinformes que no están exentos de errores de medida. Por otro lado, hay que señalar que las relaciones encontradas no implican causalidad, por lo que los resultados de este tipo de estudios de carácter exploratorio deben ir complementándose con la realización de investigaciones confirmatorias en las que se postulan hipótesis acerca de los efectos directos e indirectos existentes entre los factores implicados. En este sentido, cualquier medida educativa que se adopte para reducir la repetición, dada la relación de esta variable con el rendimiento, actuará de modo indirecto sobre este último, aunque tal aspecto habría que estudiarlo específicamente con modelos de estructuras de covarianza o con diseños experimentales o cuasiexperimentales. Como prospectiva, también sería recomendable profundizar en la replicabilidad de los resultados y comparar los de subgrupos de participantes por género, por ejemplo. Asimismo, sería interesante explorar los factores asociados a la repetición y ampliar los niveles de jerarquía para incluir al profesorado (modelo de tres niveles) e, incluso, a las comunidades autónomas (modelo con cuatro niveles).

## Contribuciones de los autores

**Alexander Constante-Amores:** Análisis formal; Curación de datos; Escritura (borrador original); Metodología.

**Deliaia Arroyo-Resino:** Análisis formal; Conceptualización; Escritura (borrador original); Escritura (revisión y edición).

**María Sánchez-Munilla:** Análisis formal; Curación de datos; Escritura (borrador original); Metodología.

**Inmaculada Asensio-Muñoz:** Conceptualización; Escritura (borrador original); Escritura (revisión y edición); Supervisión.

## Referencias bibliográficas

- Allen, C. S., Chen, Q., Willson, V. L., y Hughes J. N. (2009). Quality of design moderates effects of grade retention on achievement: A meta-analytic, multi-level analysis [La calidad del diseño modera los efectos de la retención de curso sobre el rendimiento: un análisis metaanalítico multinivel]. *Educational Evaluation and Policy Analysis*, 31 (4), 480-99. <https://doi.org/10.3102/0162373709352239>
- Arroyo, D., Constante-Amores, A., y Asensio, I. (2019). La repetición de curso a debate: un estudio empírico a partir de PISA 2015. *Educación XXI*, 22 (2), 69-92. <https://doi.org/10.5944/educxx1.22479>
- Arroyo, D., Constante-Amores, A., Castro, M., y Navarro, E. (2024a). Eficacia escolar y alto rendimiento del alumnado español en PISA 2018: un enfoque de *machine learning*. *Educación XXI*, 27 (2), 223-251. <https://doi.org/10.5944/educxx1.38634>
- Arroyo, D., Constante-Amores, A., Gil, P., y Carrillo, P. J. (2024b). Student well-being and mathematical literacy performance in PISA 2018: A machine-learning approach [Bienestar de los estudiantes y rendimiento en alfabetización matemática en PISA 2018: un enfoque de *machine learning*].

- Educational Psychology*, 44 (3), 340-357. <https://doi.org/10.1080/01443410.2024.2359104>
- Asensio, I., Carpintero, E., Expósito, E., y López, E. (2018). ¿Cuánto oro hay entre la arena? Minería de datos con los resultados de España en PISA 2015. *Revista Española de Pedagogía*, 76 (270), 225-245. <https://doi.org/10.22550/REP76-2-2018-02>
- Battistin, E., y Schizzerotto, A. (2019). Threat of grade retention, remedial education and student achievement: Evidence from upper secondary schools in Italy [La amenaza de la permanencia en el curso, la educación de recuperación y el rendimiento de los alumnos: datos de los centros de enseñanza secundaria superior en Italia]. *Empirical Economics*, 56, 651-678. <https://doi.org/10.1007/s00181-018-1443-6>
- Cameron, A. C., y Windmeijer, F. A. (1997). An R-squared measure of goodness of fit for some common nonlinear regression models [Medida R-cuadrado de la bondad de ajuste de algunos modelos habituales de regresión no lineal]. *Journal of Econometrics*, 77 (2), 329-342. [https://doi.org/10.1016/S0304-4076\(96\)01818-0](https://doi.org/10.1016/S0304-4076(96)01818-0)
- Carabaña, J. (2013). Repetición de curso y puntuaciones PISA, ¿cuál causa cuál? En Instituto Nacional de Evaluación Educativa (Ed.), *PISA 2012: Programa para la evaluación de los alumnos, informe español (volumen II: análisis secundario)*. INEE.
- Choi, A., Gil, M., Mediavilla, M., y Valbuena, J. (2018). Predictors and effects of grade repetition [Predictores y efectos de la repetición de curso]. *Revista de Economía Mundial*, (48), 21-42.
- Constante, A., Florenciano, E., Navarro, E., y Fernández, M. (2021). Factores asociados al abandono universitario. *Educación XXI*, 24 (1), 17-44. <https://doi.org/10.5944/educxx1.26889>
- Cordero, J., Manchón, C., y Simancas, R. (2014). La repetición de curso y sus factores condicionantes en España. *Revista de Educación*, (365), 12-37. <https://doi.org/10.4438/1988-592X-RE-2014-365-263>
- Ferla, J., Valcke, M., y Cai, Y. (2009). Academic self-efficacy and academic self-concept: Reconsidering structural relationships [Autoeficacia académica y autoconcepto académico: reconsiderando las relaciones estructurales]. *Learning and Individual Differences*, 19 (4), 499-505. <https://doi.org/10.1016/j.lindif.2009.05.004>

- Fernández-Lasarte, O., Ramos-Díaz, E., y Axpe, I. (2019). Rendimiento académico, apoyo social percibido e inteligencia emocional en la universidad. *European Journal of Investigation in Health*, 9 (1), 39-49.
- Gamazo, A., y Martínez-Abad, F. (2020). An exploration of factors linked to academic performance in PISA 2018 through data mining techniques [Una exploración de los factores vinculados al rendimiento académico en PISA 2018 mediante técnicas de minería de datos]. *Frontiers in Psychology*, 11. <https://doi.org/10.3389/fpsyg.2020.575167>
- García-Pérez, J., Hidalgo, M., y Robles, J. (2014). Does grade retention affect students' achievement? Some evidence from Spain [¿Influye la permanencia en un curso en el rendimiento de los alumnos? Algunos datos de España]. *Applied Economics*, 46 (12), 1373-1392. <https://doi.org/10.1080/00036846.2013.872761>
- Gaviria, J. L., y Castro, M. (2005). *Modelos jerárquicos lineales*. La Muralla.
- González-Nuevo, C., Postigo, Á., García-Cueto, E., Menéndez-Aller, Á., Muñoz, J., Cuesta, M., Álvarez-Díaz, M., y Fernández-Alonso, R. (2023). Grade retention impact on academic self-concept: a longitudinal perspective [Impacto de la permanencia en el curso sobre el autoconcepto académico: una perspectiva longitudinal]. *School Mental Health*, 15, 600-610. <https://doi.org/10.1007/s12310-023-09573-2>
- Gorostiaga, A., y Rojo-Álvarez, J. L. (2016). On the use of conventional and statistical-learning techniques for the analysis of PISA results in Spain [Sobre el uso de técnicas convencionales y de aprendizaje estadístico para el análisis de los resultados de PISA en España]. *Neurocomputing*, 171, 625-637. <https://doi.org/10.1016/j.neucom.2015.07.001>
- Hattie, J. (2017). *Aprendizaje visible para profesores. Maximizando el impacto en el aprendizaje*. Paraninfo.
- Hornstra, L., Majoor, M., y Peetsma, T. (2017). Achievement goal profiles and developments in effort and achievement in upper elementary school [Perfiles de objetivos de rendimiento y evolución del esfuerzo y el rendimiento en la escuela primaria superior]. *British Journal of Educational Psychology*, 87 (4), 606-629. <https://doi.org/10.1111/bjep.12167>
- Ikeda, M., y García, E. (2014). Grade repetition: A comparative study of academic and non-academic consequences [La repetición de curso: un estudio comparativo de las consecuencias académicas y no académicas]. *OECD Journal: Economic Studies*, (1), 269-315. [https://doi.org/10.1787/eco\\_studies-2013-5k3w65mx3hnx](https://doi.org/10.1787/eco_studies-2013-5k3w65mx3hnx)
- Jimerson, S. R. (2001). Meta-analysis of grade retention research: new directions for practice in the 21<sup>st</sup> century [Metaanálisis de la investigación sobre la retención de notas: nuevas orientaciones para la práctica en el siglo XXI]. *School Psychology Review*, 30 (3), 420-437. <http://dx.doi.org/10.1080/02796015.2001.12086124>
- Jimerson, S. R., y Ferguson, P. (2007). A longitudinal study of grade retention: Academic and behavioral outcomes of retained students through adolescence [Un estudio longitudinal de la retención escolar: resultados académicos y conductuales de los alumnos retenidos durante la adolescencia]. *School psychology quarterly*, 22 (3), 314.
- Kılıç, D., Aşkın, Ö. E., y Öz, E. (2017). Identifying the classification performances of educational data mining methods: A case study for TIMSS [Identificación de los resultados de clasificación de los métodos de extracción de datos educativos: un estudio de caso para TIMSS]. *Kuram ve Uygulamada Eğitim Bilimleri*, 17 (5), 1605-1623. <https://jestp.com/menuscrypt/index.php/estp/article/view/426>
- Kiray, S. A., Gok, B., y Bozkir, A. S. (2015). Identifying the factors affecting science and mathematics achievement using data mining methods [Identificación de los factores que afectan al rendimiento en ciencias y matemáticas mediante métodos de minería de datos]. *Journal of Education in Science, Environment and Health*, 1 (1), 28-48.
- Lee, V. E. (2000). Using hierarchical linear modeling to study social contexts: The case of school effects [Utilización de modelos lineales jerárquicos para estudiar contextos sociales: el caso de los efectos escolares]. *Educational Psychologist*, 35 (2), 125-141. [https://doi.org/10.1207/S15326985EP3502\\_6](https://doi.org/10.1207/S15326985EP3502_6)
- Liu, X., y Ruiz, M. E. (2008). Using data mining to predict K-12 students' performance on large-scale assessment items related to energy [Utilización de la minería de datos para predecir el rendimiento de los alumnos de primaria y secundaria en evaluaciones a gran escala relacionadas con la energía].

- Journal of Research in Science Teaching*, 45 (5), 554-573. <https://doi.org/10.1002/tea.20232>
- Lopes, J., Oliveira, C. y Costa, P. (2022). Determinantes escolares y de los estudiantes en el rendimiento lector: un análisis multinivel con estudiantes portugueses. *Revista de Psicodidáctica*, 27, 29-37. <http://doi.org/10.1016/j.psicod.2021.05.001>
- López, E., Expósito, E., Carpintero, E., y Asensio, I. (2018). ¿Qué nos dice PISA sobre la enseñanza y el aprendizaje de las ciencias? Una aproximación a través de árboles de decisión. *Revista de Educación*, (382), 133-161.
- López, F., y García, I. (2020). *A vueltas con la equidad en educación: una aproximación empírica en la perspectiva de las consecuencias*. Universidad Camilo José Cela. <http://hdl.handle.net/20.500.12020/901>
- López, L., González-Rodríguez, D., y Vieira, M. J. (2023). Variables que afectan la repetición en la educación obligatoria en España. *Revista Electrónica de Investigación Educativa*, 25, 1-15. <https://doi.org/10.24320/redie.2023.25.e17.4942>
- López-Rupérez, F., García-García, I., y Expósito-Casas, E. (2021). La repetición de curso y la graduación en Educación Secundaria Obligatoria en España. Análisis empíricos y recomendaciones políticas. *Revista de Educación*, (394), 325-353. <https://doi.org/10.4438/1988-592X-RE-2021-394-510>
- Marsh, H. W., Parker, P. D., y Pekrun, R. (2018). Three paradoxical effects on academic self-concept across countries, schools, and students [Tres efectos paradójicos en el autoconcepto académico de países, escuelas y estudiantes]. *European Psychologist*, 24 (3), 231-242. <http://doi.org/10.1027/1016-9040/a000332>
- Mathys, C., Véronneau, M., y Lecocq, A. (2019). Grade retention at the transition to secondary school: Using propensity score matching to identify consequences on psychosocial adjustment [Retención escolar en la transición a la enseñanza secundaria: utilización del emparejamiento de puntuaciones de propensión para identificar las consecuencias en el ajuste psicosocial]. *Journal of Early Adolescence*, 39 (1), 97-133. <https://doi.org/10.1177/0272431617735651>
- Martínez-Abad, F., y Chaparro, A. A. (2017). Data-mining techniques in detecting factors linked to academic achievement [Técnicas de minería de datos para detectar factores relacionados con el rendimiento académico]. *School Effectiveness and School Improvement*, 28 (1), 39-55. <https://doi.org/10.1080/09243453.2016.1235591>
- Martínez-Abad, F., Gamazo, A., y Rodríguez-Conde, M. J. (2020). Educational data mining: Identification of factors associated with school effectiveness in PISA assessment [Minería de datos educativos: identificación de factores asociados a la eficacia escolar en la evaluación PISA]. *Studies in Educational Evaluation*, 66, 100875. <https://doi.org/10.1016/j.stueduc.2020.100875>
- Nieto-Isidro, S., y Martínez-Abad, F. (2023). Repetición de curso y su relación con variables socioeconómicas y educativas en España. *Revista de Educación*, 1 (402), 207-236. <https://doi.org/10.4438/1988-592X-RE-2023-402-600>
- Ministerio de Educación y Formación Profesional. (2019). *PISA 2018. Programa para la evaluación internacional de los estudiantes. Informe español (vol. I)*. <https://www.educacionyfp.gob.es/inee/evaluaciones-internacionales/pisa/pisa-2018/pisa-2018-informes-es.html>
- Ministerio de Educación y Formación Profesional. (2022). *Sistema estatal de indicadores de la educación 2022*. [https://www.libreria.educacion.gob.es/libro/sistema-estatal-de-indicadores-de-la-educacion-2022\\_184171/](https://www.libreria.educacion.gob.es/libro/sistema-estatal-de-indicadores-de-la-educacion-2022_184171/)
- Niemi, H. (2015). Teacher professional development in Finland: Towards a more holistic approach [Desarrollo profesional del profesorado en Finlandia: hacia un enfoque más holístico]. *Psychology, Society and Education*, 7 (3), 279-294.
- OCDE. (2019). *PISA 2018 assessment and analytical framework [Evaluación y marco analítico de PISA 2018]*. <https://doi.org/10.1787/b25efab8-en>
- OCDE. (2020). *PISA 2018 results (volume VI). Are students ready to thrive in an interconnected world? [Resultados PISA 2018 (volumen VI) ¿Están preparados los estudiantes para prosperar en un mundo interconectado?]*. <https://www.oecd.org/publications/pisa-2018-results-volume-vi-d5f68679-en.htm>
- Pardo, M., y Ruiz, M. (2013). *Análisis de datos en ciencias sociales y de la salud III*. Síntesis.
- Peixoto, F., Monteiro, V., Mata, L., Sanches, C., Pipa, J., y Navas, L. (2016). *To be or not to be retained... That's the question! Retention, self-esteem, self-concept, achievement goals, and grades [Ser o no ser retenido... ¡Esa es la cuestión! Retención, autoestima, autoconcep-*

- to, objetivos de rendimiento y calificaciones]. *Frontiers in Psychology*, 7, (1550). <https://doi.org/10.3389/fpsyg.2016.01550>
- Raschka, S. (2015). *Python machine learning [Aprendizaje automático en Python]*. Packt.
- Rhodes, J., Thomas, J. M., y Liles, A. R. (2018). Predictors of grade retention among children in an elementary school truancy intervention [Factores que predicen la permanencia en el colegio de los niños que participan en una intervención contra el absentismo escolar en la escuela primaria]. *Journal of At-Risk Issues*, 21 (1), 1-10.
- Rivero, G. (2011). *Análisis de datos incompletos en ciencias sociales*. CIS.
- Rodríguez-Rodríguez, D. (2022). Repetición de curso, rendimiento académico y variables motivacionales en Educación Secundaria Obligatoria: un estudio longitudinal. *Psicothema*, 34 (3), 429-436. <https://doi.org/10.7334/psicothema2021.582>
- Rosenthal, R., y Rosnow, R. L. (2008). *Essentials of behavioral research: Methods and data analysis [Fundamentos de la investigación conductual: métodos y análisis de datos]*. McGraw-Hill
- Sarkar, D., Bali, R., y Sharma, T. (2018). *Practical machine learning with Python. A problem-solvers guide to building real-world intelligent systems [Aprendizaje automático práctico con Python. Guía de resolución de problemas para crear sistemas inteligentes reales]*. Apress
- Schwerdt, G., West, M. R., y Winters, M. A. (2017). The effects of testbased retention on student outcomes over time: Regression discontinuity evidence from Florida [Los efectos de la retención basada en exámenes sobre los resultados de los estudiantes a lo largo del tiempo: pruebas de regresión discontinua en Florida]. *Journal of Public Economics*, 152, 154-169. <https://doi.org/10.1016/j.jpubeco.2017.06.004>
- Seabra, D. A., y Ferrão, M. E. (2016). Repetência e indisciplina: Evidências de Brasil e Portugal no PISA 2012 [Repetición e indisciplina: datos de Brasil y Portugal en PISA 2012]. *Cadernos de Pesquisa*, 46 (161), 614-636. <https://publicacoes.fcc.org.br/cp/article/view/3669>
- Tapia, J. G., y Álvarez, C. A. (2022). La repetición de curso en educación secundaria en clave organizacional. *Avances en Supervisión Educativa*, (38). <https://doi.org/10.23824/ase.v0i38.772>
- Urbina, A. B., y De la Calleja, J. (2017). Breve revisión de aplicaciones educativas utilizando minería de datos y aprendizaje automático. *Revista Electrónica de Investigación Educativa*, 19 (4), 84-96. <https://doi.org/10.24320/redie.2017.19.4.1305>
- Valbuena, J., Mediavilla, M., Choi, Á., y Gil, M. (2021). Effects of grade retention policies: A literature review of empirical studies applying causal inference [Efectos de las políticas de retención de notas: una revisión bibliográfica de estudios empíricos que aplican la inferencia causal]. *Journal of Economic Surveys*, 35 (2), 408-451. <https://doi.org/10.1111/joes.12406>
- Van Canegem, T., Van Houtte, M., y Demanet, J. (2021). Grade retention and academic self-concept: A multilevel analysis of the effects of schools' retention composition [Retención escolar y autoconcepto académico: un análisis multinivel de los efectos de la composición de la retención escolar]. *British Educational Research Journal*, 47 (5), 1340-1360. <https://doi.org/10.1002/berj.3729>
- Warren, J. R., Hoffman, E., y Andrew, M. (2014). Patterns and trends in grade retention rates in the United States, 1995-2010 [Patrones y tendencias de las tasas de retención escolar en Estados Unidos, 1995-2010]. *Educational Researcher*, 43 (9), 433-443. <https://dx.doi.org/10.3102/0013189X14563599>

## Biografía de los autores

**Alexander Constante-Amores.** Graduado en Pedagogía y Máster en Investigación en Educación en la Universidad Complutense de Madrid (UCM). En la actualidad, está finalizando el Programa de Doctorado en Educación en la UCM. Es profesor de la Universidad Camilo José Cela y de la Universidad Europea de Madrid, donde imparte asignaturas en las áreas de Estadística, Bioestadística y Métodos de Investigación. Ha sido becario del Instituto Nacional de Estadística y es miembro del grupo de investigación Medida y Evaluación de Sistemas Educativos (MESE).

Su línea de investigación se centra en evaluación de sistemas educativos.

 <https://orcid.org/0000-0003-2024-9615>

**Delia Arroyo-Resino.** Profesora contratada doctora en la Universidad Complutense de Madrid. Doctorado Internacional con premio extraordinario en el Departamento de Métodos de Investigación y Diagnóstico en Educación por la Universidad Complutense de Madrid. Miembro del grupo de investigación de Medida y Evaluación de Sistemas Educativos.

 <https://orcid.org/0000-0002-3784-7745>

**María Sánchez-Munilla.** Graduada en Pedagogía y Máster en Metodologías de las Ciencias del Comportamiento de la Salud en la Universidad Nacional de Educación a Distancia (UNED). En la actualidad, está finalizando el Programa de Doctorado en Educación en la UCM con un contrato predoctoral de Formación del Profesorado Universitario (FPU). Es miembro del grupo de investigación Medida y Evaluación de Sistemas Educativos (MESE) y del Servicio de Evaluación y Diagnóstico en

Educación (SEDE) de la Facultad de Educación de la UCM.

 <https://orcid.org/0000-0001-6689-7170>

**Inmaculada Asensio-Muñoz.** Doctora con premio extraordinario por la Universidad Complutense de Madrid. Profesora titular del Área de Métodos de Investigación y Diagnóstico en Educación, en el Departamento de Investigación y Psicología en Educación de la Facultad de Educación en dicha Universidad. Tiene una amplia experiencia docente en pedagogía y en formación del profesorado, tanto a nivel de grado como de máster y doctorado. Es miembro del grupo de investigación Medida y Evaluación de Sistemas Educativos (MESE) y, como experta en metodología de investigación educativa, en su trayectoria profesional, ha trabajado en diversos proyectos de investigación e innovación financiados relacionados con esa temática. Sus publicaciones se centran en la mejora de la enseñanza y del aprendizaje y, en general, de la función docente y orientadora.

 <https://orcid.org/0000-0003-3523-570X>