

DetECCIÓN TEMPRANA DEL RIESGO ESCOLAR. Predicción de trayectorias de rezago en la educación primaria en Uruguay mediante técnicas de *machine learning* Early Detection of School Risk. Prediction of Trajectories of School Backwardness in Primary Education in Uruguay Using Machine Learning Techniques

Santiago Cardozo
UNIVERSIDAD DE LA REPÚBLICA, URUGUAY
cardozo.santiago@gmail.com
ORCID 0000-0002-7767-1617

Adrián Silveira
UNIVERSIDAD DE LA REPÚBLICA, URUGUAY
adrian.silveira@fic.edu.uy
ORCID 0000-0002-1899-3368

Bruno Fonseca
ANEP-CODICEN, URUGUAY
bfonsecalema@gmail.com
ORCID 0000-0002-2327-9695

RESUMEN

El trabajo utiliza técnicas de aprendizaje computacional (*machine learning*) para la estimación del riesgo escolar durante los primeros tres años de la trayectoria en la enseñanza primaria de una cohorte de alumnos uruguayos. Se utilizan tres técnicas de análisis (regresión logística, redes bayesianas y árboles de clasificación) con el propósito de identificar el riesgo de trayectorias escolares pautadas por la repetición de al menos un curso, en función de un conjunto de factores antecedentes a su transición al primer grado de enseñanza primaria. Estos factores abarcan desde las condiciones sociosanitarias al momento del nacimiento hasta la situación familiar y educativa de los alumnos sobre el final de su escolarización preprimaria. En particular, el análisis se focaliza en el poder predictivo de las habilidades captadas por la Evaluación Infantil Temprana (EIT) que se aplica próximo a la finalización de la educación inicial, sobre los cinco años de edad. Los resultados sugieren que las habilidades captadas por EIT logran identificar anticipadamente a la mayoría de los niños con riesgo educativo. Los niveles de precisión y de sensibilidad de los modelos que incluyen este factor evidencian el potencial de los sistemas de alerta temprana para detectar y prevenir situaciones de "fracaso escolar".

Palabras clave: desarrollo infantil, riesgo escolar, machine learning, repetición

ABSTRACT

The work uses computational learning techniques (*machine learning*) to estimate school risk during the first three years of the trajectory in primary education of a cohort of Uruguayan students. We use three analysis techniques (logistic regression, Bayesian networks, and classification trees) to identify the risk of school trajectories based on the repetition of at least one year, based on a set of factors before the transition to the first year of primary school. These factors range from the students' socio-sanitary conditions at their birth to their family and educational situation at the end of their pre-primary schooling. In particular, the analysis focuses on the predictive power of the skills captured by the Early Childhood Assessment (EIT) that is applied close to the end of initial education, around five years of age. The results suggest that the skills captured by EIT manage to identify in advance the majority of children with educational risk. The levels of precision and sensitivity of the models that include this factor show the potential of early warning systems to detect and prevent situations of "school failure".

Keywords: child development, school risk, machine learning, repetition

INTRODUCCIÓN

Los niños comienzan su escolarización en condiciones desiguales para enfrentar las demandas cognitivas y comportamentales que requiere la educación formal y para aprovechar las oportunidades de aprendizaje que ofrece la escuela. Parte importante de estas desigualdades se expresa en el nivel de desarrollo de un conjunto de habilidades cognitivas, lingüísticas, socioemocionales y físicas, y en una serie de predisposiciones y actitudes hacia los aprendizajes, que la literatura especializada denomina genéricamente con el término *Preparación Para la Escuela* (PPE). La PPE depende de los procesos de maduración y, en consecuencia, de la edad, pero también del tipo de interacción del niño con diversos factores ambientales, familiares e institucionales, desde la concepción, el nacimiento y durante la infancia. La investigación educativa ha mostrado que las habilidades vinculadas a la PPE son altamente predictivas de la trayectoria escolar posterior; en particular, que un insuficiente nivel de desarrollo temprano en estas dimensiones supone importantes riesgos para la escolarización.

Este artículo propone evaluar la capacidad predictiva de un conjunto habilidades vinculadas a la PPE, captadas al término de la educación inicial, sobre los riesgos de repetir algún curso durante los primeros tres años en la enseñanza primaria. Se utiliza información del Panel-EIT, un estudio longitudinal basado en el seguimiento de las trayectorias educativas de los niños uruguayos que, sobre los cinco años de edad, participaron de la Evaluación Infantil Temprana (EIT) en 2016.

El artículo busca contribuir en tres niveles diferentes. En primer lugar, mostrar que un instrumento como EIT, aplicado por los propios docentes en contextos escolares cotidianos y fácilmente escalable al conjunto de la matrícula, puede aportar insumos robustos sobre los cuales basar criterios de clasificación de alumnos según su riesgo escolar, de cara a la transición a la escuela graduada.

En términos metodológicos, en segundo lugar, el trabajo hace uso de técnicas de aprendizaje computacional (*machine learning*), especialmente útiles para aprovechar la creciente disponibilidad de información con la que cuentan actualmente los sistemas educati-

vos. Estos modelos permiten aprender criterios de clasificación en función de los riesgos educativos futuros de los alumnos, con base en el análisis intensivo de datos comparativamente grandes. Aunque hacen uso de técnicas estadísticas relativamente complejas, los resultados de este proceso pueden ser traducidos a protocolos simples de clasificación, fácilmente comunicables a la comunidad educativa, docentes y a las familias.

Tercero, estos análisis pretenden contribuir al diseño de políticas educativas en tres aspectos complementarios: i) mostrando la potencialidad de definir estrategias de intervención en función de la identificación de los riesgos educativos individuales, como alternativa o complemento a los criterios de focalización predominantes en la mayoría de los sistemas educativos de la región, basados en la escuela como unidad de intervención; ii) el énfasis en la detección temprana y en la noción de riesgo, entendido como la anticipación de una trayectoria probablemente vulnerable, por oposición al desarrollo de estrategias remediales que operan después que el fracaso escolar se manifestó y que, por tanto, pueden resultar tanto ineficientes como inoportunas; iii) la identificación de los riesgos con base en las propias habilidades que afectan las oportunidades de éxito en la escuela y que son pasibles de una intervención pedagógica específica, más que en variables de tipo contextual (como el entorno socioeconómico de la escuela, las características de las familias o la localización geográfica) que inciden en los logros educativos predominantemente de manera indirecta.

A su vez, Uruguay constituye un caso de especial interés para valorar la potencia de un sistema de identificación temprana de riesgos escolares basado en la medición de las habilidades vinculadas a la PPE: la universalización, de forma pionera en la región, del acceso de los niños de cuatro y cinco años a la educación inicial posibilita la evaluación a gran escala en las propias escuelas; los potentes sistemas de información administrativa a nivel de alumnos (GURI-CEIP),¹ con una importante acumulación en el diseño e implementación de evaluaciones a gran escala para alumnos de nivel inicial (EIT e

¹ Es el sistema de información web del Consejo de Educación Inicial y Primaria (CEIP) de la Administración Nacional de Educación Pública (ANEP) de Uruguay.

INDI)² y con sistemas universales de evaluación formativa durante la enseñanza primaria (SEA).³

Asimismo, desde las políticas educativas, Uruguay ha definido en los últimos años una batería de acciones orientadas hacia la intervención oportuna y personalizada de niños en situación de riesgo escolar, entre las que se destacan: el Protocolo de Prevención del Ausentismo, el Protocolo de Seguimiento de niños de primer grado y el Plan de Trabajo Personal (PTP), de carácter prescriptivo, para alumnos de primaria que repiten u obtienen fallos de promoción extraordinaria, así como mecanismos de alerta vinculados a la gestión local y a tiempo real, de situaciones de ausentismo a clases. A la fecha, estas orientaciones no han derivado, sin embargo, en un sistema consolidado de alertas tempranas de alcance universal con base en estudios sistemáticos de predicción de riesgos educativos.

REVISIÓN DE LITERATURA

El desarrollo infantil y la preparación para la escuela

La preparación para la escuela (PPE) refiere al nivel de desarrollo de un conjunto de habilidades o competencias básicas consideradas prerrequisitos para insertarse exitosamente en la escuela, tanto en términos académicos como sociales (UNESCO, 2018; UNICEF, 2012; Raver *et al.*, 2011; Duncan *et al.*, 2007; Hair *et al.*, 2006). Por lo general, estas habilidades se valoran al inicio de la escolarización formal, ya sea al ingreso al nivel inicial o durante la transición a la escuela graduada, entre el nivel inicial y la educación primaria.

²EIT es la Evaluación Infantil Temprana aplicada en Uruguay hasta 2017 a niños de cuatro y cinco años de inicial de escuelas públicas. Valora el nivel de desarrollo en cinco dimensiones: conciencia de sí mismo y del entorno, habilidades sociales y enfoques para el aprendizaje, habilidades cognitivas, lenguaje y comunicación y desarrollo físico. EIT es una adaptación a Uruguay del *Early Years Evaluation* (Sloat, Beswick y Willms, 2007). INDI es el Inventario de Desarrollo Infantil (Vásquez y Moreira, 2016), el instrumento mediante el que actualmente se evalúa la disposición para la escolarización de los niños de los niveles 3, 4 y 5 de educación inicial en cuatro áreas del desarrollo: cognitiva, motora, socioemocional y vinculada a la disposición para el aprendizaje.

³El Sistema de Evaluación de Aprendizajes (SEA) es una plataforma de evaluaciones en línea de carácter formativo desarrollado e implementado por la ANEP.

La PPE es esencialmente multidimensional (Vázquez y Moreira, 2016): involucra diversos dominios vinculados al desarrollo –cognitivo, lingüístico, socioemocional, físico-motor–, así como al bienestar sanitario, junto con un conjunto de actitudes y predisposiciones hacia el aprendizaje, asociados, entre otros aspectos, al interés y a la curiosidad frente a nuevas tareas o a la capacidad de persistencia y atención, entre otras. En general, se asume que cada una de estas dimensiones es, por sí misma, una condición necesaria pero no suficiente para una adecuada PPE, aunque no existen consensos definitivos en torno a su importancia relativa (Allotey *et al.*, 2017; Wilson, 2014; La Paro y Pianta, 2000).

Empíricamente, la investigación muestra que la PPE no es sólo el resultado de los procesos naturales de maduración asociados a la edad, sino que depende fuertemente de la interacción, a lo largo de la infancia, con diversos factores ambientales, familiares e institucionales (López y Willms, 2020; Rebello, 2012; Wilson, 2014; Lee y Burkam, 2002). En consecuencia, las desigualdades en la PPE están estrechamente asociadas, entre otros factores, a dimensiones estructurales como el nivel socioeconómico y educativo de la familia de origen (Lee y Burkam, 2002), la condición étnico-racial y lingüística, las condiciones sanitarias y sociales en el nacimiento y durante la primera infancia (Amarante y Labat, 2018; Allotey *et al.*, 2017; Luu *et al.*, 2017; Almond y Currie, 2010), así como a variables demográficas típicamente exógenas, como el sexo y la edad (Hair *et al.*, 2006; Gledhill *et al.*, 2002; Arrivillaga *et al.*, 2016; Vázquez y Moreira, 2016).

En síntesis, bajo el enfoque que asume este artículo, la PPE engloba un conjunto de aspectos vinculados al desarrollo cognitivo, socioemocional y físico de los niños al momento de su transición a la escuela graduada, conceptualizados como condiciones necesarias –aunque no suficientes– para una trayectoria escolar exitosa, en términos académicos y sociales. El foco de nuestra preocupación radica en la constatación de importantes brechas en las condiciones con la que los niños inician su escolarización y, más específico, en la necesidad de definir protocolos sencillos, pero de alta sensibilidad, que permitan identificar de manera temprana a aquellos que, con alta probabilidad, enfrentarán riesgos serios para progresar en la es-

cuela y para desarrollar los aprendizajes previstos en cada etapa. Esto implica que las escuelas deberían lograr incorporar la diversidad de situaciones de partida en su planificación y en sus rutinas cotidianas, de modo de incluir efectiva y eficazmente a todos los niños en las aulas; en otros términos, las escuelas deben estar “preparadas para los niños”, tanto como los niños necesitan estar preparados para la escuela (Bhattacharjea, 2019).

La preparación para la escuela y el logro escolar

Desde la década de los noventa existe una acumulación sistemática de evidencia sobre el poder predictivo de las habilidades vinculadas a la PPE respecto al desempeño educativo posterior, ya sea vinculado al logro específicamente académico, ya en términos de la integración o el involucramiento con la escuela, entre otros (López y Willms, 2020; Linan-Thompson, 2014, Wilson, 2014, National Early Literacy Panel, 2008). Entre los resultados considerados por los estudios sobre el tema, se encuentran indicadores vinculados a la repetición o progresión en la trayectoria (Davoudzadeh *et al.*, 2015; Duncan *et al.*, 2007; Blair, 2001), así como el desempeño en evaluaciones estandarizadas de conocimiento general o en áreas específicas como matemáticas y lenguaje (Pan *et al.*, 2019; Jung, 2016; Christopolou y Pampaka, 2015; Clark *et al.*, 2010; Scarborough, 2001). Algunos estudios se han ocupado, asimismo, del poder predictivo de las habilidades vinculadas a la PPE sobre otras dimensiones no específicamente educativas, por ejemplo, de tipo sanitario (Hair *et al.*, 2006).

La importancia relativa de los diversos dominios de la PPE para la predicción del logro académico posterior es un asunto de controversia entre los especialistas. En general, existe consenso en torno al poder predictivo de las llamadas habilidades preacadémicas, vinculadas al desarrollo cognitivo y lingüístico. Diversos metaanálisis sugieren, en este sentido, que las habilidades tempranas en estos dominios tienen un poder predictivo sustantivamente mayor en comparación con otras dimensiones más “blandas”, como las habilidades socioemocionales, actitudinales o físicas (Wilson, 2014; Duncan *et al.*, 2007). De todos modos, los estudios sugieren que el peso relativo de las diferentes dimensiones de la PPE depende, entre

otros aspectos, del indicador de logro considerado (Rimm-Kaufman y Sandilos, 2017).

LA INVESTIGACIÓN SOBRE EL DESARROLLO INFANTIL Y EL DESEMPEÑO ESCOLAR EN URUGUAY

En Uruguay existe una creciente acumulación de investigación vinculada al desarrollo infantil y su asociación con las brechas en los logros educativos posteriores. El estudio pionero de Mara *et al.* (2000) evaluó el nivel de desarrollo lingüístico y cognitivo en niños de cuatro y cinco años y mostró su asociación con los riesgos de repetición posterior en primaria. Los autores encuentran, además, que el nivel de desarrollo a los cuatro años dependía del contexto sociofamiliar, pero también del grado de exposición a la educación inicial y del tipo de prácticas pedagógicas de los maestros en este nivel.

Más recientemente los trabajos de Failache, Salas y Vigorito (2015), Cabella *et al.*, (2015) y de Failache y Katzkowicz (2018), con base en la Encuesta de Nutrición, Desarrollo Infantil y Salud (ENDIS), reportaron asociaciones significativas, por un lado, entre las condiciones de vida en la primera infancia y el desarrollo infantil en diferentes dominios, y, por otro, entre el nivel de desarrollo infantil y distintos indicadores de logro educativo. A similares conclusiones llegan Borba *et al.* (2018), respecto a los factores determinantes de las habilidades captadas por la Evaluación Infantil Temprana y a la asociación entre estas habilidades y los riesgos de repetición en primer año. Liz (2018), por su parte, reporta asociaciones significativas entre la subescala de funcionamiento ejecutivo, relevada por el INDI en el nivel 5 de educación inicial, y el involucramiento escolar de los alumnos en primer año de primaria. Por último, López y Willms (2020) utilizaron la evaluación EIT para analizar el impacto de un año de exposición a la educación inicial sobre el desarrollo cognitivo y lingüístico de los niños uruguayos de cuatro años, más allá de los efectos de “maduración” propios de la edad.

DATOS Y MÉTODOS

Fuentes de información

Los datos provienen del estudio longitudinal de trayectorias educativas de la cohorte participante de EIT en 2016 (en adelante, Panel-EIT). El panel se conformó con base al seguimiento de la cohorte de estudiantes de nivel 5 de Inicial que participaron de EIT en 2016 y comenzaron el primer grado de educación primaria en 2017 (n = 15.529).

El Panel-EIT articula información sobre diversas dimensiones, con base en fuentes complementarias: i) información perinatal sobre las condiciones de la gestación, el embarazo y el parto, derivada de los registros del Carnet de Nacido Vivo (CNV) del Ministerio de Salud Pública de Uruguay; ii) información sobre el contexto socioeconómico de las familias (hogares beneficiarios de los programas Tarjeta Uruguay Social y Asignaciones Familiares del Plan de Equidad) del Ministerio de Desarrollo Social de Uruguay; iii) información sociodemográfica y sobre la trayectoria escolar, incluidos los fallos de promoción/repetición, de los estudiantes y de las escuelas públicas de Uruguay; vi) resultados del estudio EIT en 2016.

La Evaluación Infantil Temprana (EIT)

EIT evalúa cinco dominios vinculados a la preparación para la escuela (cuadro 1). Es aplicado por los maestros en el aula y ofrece un marco de prevención asociado a la elaboración de protocolos de intervención por “niveles de respuesta” (López y Salsamendi, 2018).

EIT clasifica a cada niño en tres niveles: i) *verde* (el niño logra realizar las tareas correspondientes sin dificultades); ii) *amarillo* (el niño experimenta ciertas dificultades), y iii) *rojo* (presenta dificultades severas o significativas). Estas tres categorías se repiten para cada uno de los dominios evaluados y para el indicador global que los resume (Willms *et al.*, 2000).

■ Cuadro 1. Dimensiones evaluadas por la Evaluación Infantil Temprana 2016

| Dimensión o dominio | Qué mide | Núm. de ítems |
|--|--|---------------|
| Conciencia de sí mismo y del entorno (CME) | La comprensión que tiene el niño del mundo y su capacidad para relacionarse con las experiencias en el hogar y en la comunidad. | 8 |
| Habilidades sociales y enfoques para el aprendizaje (HSEA) | Alude a la atención que pone el niño en las actividades en el salón de clase y su capacidad para interactuar con sus compañeros al mismo tiempo que respeta las reglas de salón. | 8 |
| Habilidades cognitivas (HC) | Se relaciona con las habilidades básicas del niño en matemáticas y lectura, y su capacidad para resolver problemas. | 8 |
| Lenguaje y comunicación (LC) | Se refiere a la comprensión del niño del lenguaje hablado y su capacidad para expresar pensamientos y sentimientos. | 8 |
| Desarrollo físico(MF+MG) | a) Motricidad fina: es la capacidad del niño de realizar pequeños movimientos que requieren la coordinación mano-ojo; b) Motricidad gruesa: refiere a la capacidad del niño para realizar movimientos amplios que involucren brazos, piernas y cuerpo. | 10 (5+5) |

Fuente: elaboración propia con base en Willms *et al.* (2000).

Este artículo consideró una medida alternativa de habilidad global desarrollada en el marco del Panel-EIT, presentada en Borba *et al.* (2018). En específico, la medida de desempeño global diseñada para el Panel-EIT se define a partir de las siguientes reglas: i) *verde* cuando todos los dominios son verdes y hasta uno es clasificado como *amarillo*; amarillo teniendo dos o más dominios amarillos y ninguno rojo; y *rojo* cuando el estudiante presenta al menos un dominio en rojo.

Métodos de estimación del riesgo escolar

El enfoque metodológico asumido en este artículo se focaliza en el uso intensivo de datos para la predicción de un resultado de interés –en este caso, una trayectoria escolar impactada por al menos un evento de repetición–. El propósito estriba en definir los criterios para clasificar a cada niño, con los mayores niveles de precisión, especificidad y sensibilidad posibles, según esté expuesto o no a una situación sustantiva de “riesgo escolar” *al inicio de su trayectoria*, es decir, antes de que el “fracaso” se manifieste.

En particular, se estimaron tres modelos de riesgo escolar mediante técnicas de aprendizaje computacional supervisado y técnicas de remuestreo. En específico, se evaluó la capacidad de los modelos para predecir trayectorias de rezago durante los primeros tres años en la enseñanza primaria, pautadas por al menos un evento de repetición del curso, con base en las habilidades valoradas por EIT a los cinco años de edad y a una batería amplia de indicadores sociodemográficos y escolares antecedentes a la transición a la escuela graduada.

El primer modelo corresponde a una regresión logística, con base en la siguiente especificación:

$$(1) \quad q = \hat{p}(y) = G(\beta_0 + x\beta) = G(Z)$$

$$(2) \quad G(Z) = \frac{e^{(Z)}}{1 + e^{(Z)}}$$

donde y es una variable binaria que asume el valor 1 si el alumno experimentó al menos un evento de repetición entre 2017 y 2019 y corresponde al conjunto de variables predictivas.

El segundo método se basa en un modelo de *Redes Bayesianas* (BN). Una BN es un modelo probabilístico que representa las dependencias condicionales entre variables aleatorias, a través de un Gráfico Acíclico Dirigido (GAD). El GAD se forma a partir de nodos, que corresponden a las variables (x_1, x_2, \dots, x_k) y de arcos que representan las dependencias entre ellas. $Pa(x_i)$ denota al conjunto de padres de X_i , tal que Y es un padre de X si y sólo si existe en la red un arco $Y \rightarrow X$.

Cada nodo tiene una función de probabilidad asociada que toma como *input* exclusivamente los valores de sus variables padres, es decir, $p(x_i | Pa(x_i))$, o su probabilidad marginal para las variables que no tienen padres. La estructura de la red puede ser definida de acuerdo con un modelo conceptual de interacción entre las variables o aprendida a partir de los datos. Se optó por un camino intermedio, definiendo algunas restricciones (duplas de variables para las cuales no hace sentido una dependencia directa) y permitiendo que el resto de las interacciones sean aprendidas a partir de los datos. En específico, se utilizó el método *Hill-climbing*, basado en tres operaciones iterativas de búsqueda local (agregar, eliminar o cambiar la dirección de un arco). El algoritmo itera en torno a las diferentes

estructuras posibles para representar la red y selecciona aquella que maximiza el ajuste, es decir, la verosimilitud del resultado. Se adoptó el criterio de ajuste BIC que penaliza la solución final en función de su complejidad (número de casos y parámetros), tal que:

$$(3) \text{BIC} = \log(p(D|\hat{\theta}, G)) - \frac{n_p}{2} \log(N),$$

$$p(D|\hat{\theta}, G)$$

donde $p(D|\hat{\theta}, G)$ es el máximo valor de la función de verosimilitud de la red G , $\hat{\theta}$ son los valores de los parámetros (probabilidades condicionadas) que maximizan la función de verosimilitud, N es el número de casos y n_p es el número de parámetros.

Una vez aprendida la estructura de la red las probabilidades condicionadas $p(x_i|Pa(x_i))$ se estiman de forma de maximizar la función de verosimilitud conjunta, definida por:

$$(4) p(x_1, x_2, \dots, x_k) = \prod_{i=1}^k p(x_i | Pa(x_i))$$

El tercer modelo corresponde a un *árbol de clasificación*, un método predictivo tipo CART (*classification and regression trees*) utilizado para el caso de variables de resultado discretas. Los CART se basan en particiones sucesivas de las observaciones en subclases (nodos, ramas y hojas), tomando una variable predictiva a la vez como criterio de clasificación, hasta alcanzar un umbral determinado de ajuste. La idea fundamental es que los nodos “hijos” sean más puros que los nodos “padres” con respecto a la clasificación de interés.

El algoritmo utilizado, *rpart* (*recursive partitioning and regression trees*), selecciona para cada partición la variable predictora y el punto de corte, si corresponde, que logra la menor heterogeneidad o impureza promedio entre las dos subparticiones resultantes. La medida de impureza utilizada es el Coeficiente de Gini que, para una variable de resultado binaria, viene dado por:

$$(5) i_G(t)_j = 2p(1-p)$$

evaluado en cada nodo t , donde p es la proporción de casos que toman valor 1 para la variable de respuesta en cada clase j ($J=2$).

El algoritmo de clasificación, parte de la consideración de todas las variables y se detiene cuando: a) una nueva partición no supone

una reducción significativa en $i_G(t)_j$; o; b) cuando una nueva partición da lugar a un nodo más pequeño que cierto umbral definido *a priori* (en este caso, un nodo con menos del 0.5% de la muestra).

Criterios de valoración

Para evaluar la capacidad predictiva de los modelos se utilizó la técnica de validación cruzada *leave-one-out cross validation*. El procedimiento consiste en entrenar el modelo empleando todas las observaciones disponibles excepto una, que se excluye para emplearla como test. El proceso se repite tantas veces como observaciones disponibles, excluyendo en cada iteración una observación distinta y ajustando el modelo con el resto. Tras iteraciones, se obtiene la probabilidad estimada

$$q = \hat{p}(Y=1|x_1, x_2, \dots, x_k), \{0 \leq q \leq 1\}$$

para la totalidad de la muestra, siendo Y la variable de respuesta y x_1, x_2, \dots, x_k la realización de las variables predictoras.

En este caso q es la probabilidad estimada de que el estudiante repita al menos una vez en los primeros tres años en primaria A efectos de la clasificación, es necesario definir un umbral, π , a partir del cual el alumno se considerará “en riesgo”. Esta definición es crítica: valores bajos π implican una mayor sensibilidad del modelo, es decir, una mejor capacidad para captar a los estudiantes “verdaderamente” en riesgo, pero la sensibilidad se gana a costa de una pérdida de especificidad. Inversamente, un valor alto de π supone una mejora en la especificidad, pero implica una pérdida de sensibilidad para detectar poblaciones que deberían ser intervenidos.

En otras palabras, existe un *trade-off* entre la sensibilidad y la especificidad: la mejora en la capacidad para identificar estudiantes en riesgo suele tener, como contrapartida, un costo en términos de la especificidad de la solución y viceversa.⁴ Una solución frecuen-

⁴El AUC (*Area Under the ROC Curve*) es un estimador del *trade-off* entre la sensibilidad y la especificidad del modelo, basado en el cálculo del área bajo la curva ROC (*Receiver Operating Characteristic Curve*), que representa la proporción de VP y de FP para un ajuste determinado en un eje cartesiano bidimensional. Constituye un indicador clásico de ajuste de modelos predictivos y una herramienta visualmente útil cuando las q estimadas siguen una distribución aproximadamente continua. Dado que en este artículo sólo aplica a la especificación logística, no reportamos AUC entre los indicadores de ajuste.

te consiste en establecer π a partir de la proporción que experimenta el evento de interés en el conjunto de la población o de la muestra (En este caso: $\pi = 0.16$). Este artículo definió $\pi = 0.20$, un valor cercano pero superior a la probabilidad marginal de experimentar un evento de repetición, que, de acuerdo con las pruebas realizadas, maximiza el equilibrio entre sensibilidad y especificidad.⁵

La valoración del desempeño de los tres modelos presentados se basa en la matriz de confusión (cuadro 2), que surge de aplicar el umbral π a la probabilidad estimada en cada caso y que permite clasificar a cada niño de la muestra como: verdadero positivo (VP), verdadero negativo (VN), falso positivo (FP) y falso negativo (FN).

■ Cuadro 2. Matriz de confusión para el evento de repetición

| Observado | Predicho | |
|-----------|-----------|--------|
| | No repite | Repite |
| No repite | VN | FP |
| Repite | FN | VP |

En virtud de esta clasificación, definimos los siguientes indicadores de ajuste:

$$(i) \textit{ Precisión (accuracy)} = \frac{(VP)+(VN)}{(VP)+(VN)+(FP)+(FN)}$$

corresponde a la proporción global de casos correctamente clasificados por el modelo. Su inverso, $1 - \textit{accuracy}$, es la tasa de error.

$$(ii) \textit{ Kappa} = \frac{Pr(a)-Pr(e)}{1-Pr(e)}$$

es un ajuste de (i) en función de la probabilidad de una predicción correcta basada exclusivamente en el azar. En (ii), $Pr(a)$ es la proporción de casos correctamente predichos por el modelo y $Pr(e)$ es la proporción esperada de aciertos que se obtendrían al azar dadas las distribuciones marginales de la matriz, a la manera de un test clásico de χ^2 .

⁵ De todos modos, en un escenario de aplicación práctica, no existen impedimentos para utilizar más de un umbral para definir intervenciones cualitativa o cuantitativamente diferenciadas.

$$(iii) \text{ Sensibilidad} = \frac{(VP)}{(VP)+(FN)}$$

es la proporción de casos positivos correctamente identificados por el modelo. Aquí, corresponde a la capacidad para detectar a los niños que efectivamente terminarán repitiendo algún grado escolar en los años posteriores.

$$(iv) \text{ Especificidad} = \frac{(VN)}{(VN)+(FP)}$$

es la proporción de casos negativos correctamente identificados por el modelo. Corresponde a la capacidad del ajuste para no incluir innecesariamente casos en el set de alumnos en riesgo, cuando en realidad no lo están.

RESULTADOS

Organizamos los resultados en tres secciones. En la primera, describimos los riesgos acumulados de repetición que enfrentó el Panel-EIT en los primeros tres años en la enseñanza primaria. En la segunda, reportamos los resultados de los tres modelos predictivos (regresión, redes bayesianas y árboles de clasificación), mientras que en la última valoramos su ajuste en términos de los criterios expuestos en la sección anterior.

Riesgos de repetición en los primeros tres años en la enseñanza primaria

El 16.2% de los alumnos del Panel-EIT repitió al menos una vez durante la ventana de observación (cuadro 3). Este valor, que corresponde a la probabilidad marginal de evento de interés, puede ser interpretado como una función acumulada de los riesgos específicos en cada año t en primaria, a partir de:

$$(1) q_{t,t+1} = P(t < T < t + 1 | T < t)$$

es decir, la probabilidad de que el alumno repita en cada año t ($t=1, 2, 3$), dado que no repitió antes.

$$(2) S_t = Pr(T \geq t) = S_{t-1} * p_{t,t+1} \quad (S_0 = 1)$$

es la función de supervivencia: indica la probabilidad de que un estudiante no haya repetido antes del año t .

(3) $1 - S_t$

es el complemento de (2): informa sobre la proporción acumulada de alumnos con al menos un evento de repetición al término de cada año t .

Como se observa en el cuadro 3, los riesgos de repetición ($q_{t, t+1}$) se concentran fuertemente en el año de ingreso a la escuela graduada ($q_{t=1}=0.114$) y decrecen de manera progresiva en los dos años posteriores ($q_{t=2} = 0.033$ y $q_{t=3} = 0.021$).

■ Cuadro 3. Tabla de supervivencia para el evento de repetición 2017-2019

| T | $N_t^{(*)}$ | $d_t^{(**)}$ | $q_{t, t+1}$ | S_t | $1-S_t$ |
|------------|-------------|--------------|--------------|-------|---------|
| t=1 (2017) | 15,529 | 1,775 | 0.114 | 0.886 | 0.114 |
| t=2 (2018) | 13,754 | 459 | 0.033 | 0.856 | 0.144 |
| t=3 (2019) | 13,295 | 278 | 0.021 | 0.838 | 0.162 |

Fuente: elaboración propia en base al Panel-EIT (*) Corresponde al conjunto en riesgo en cada año t , es decir, al número de estudiantes que no habían experimentado un evento de repetición al inicio de cada periodo. (**) Corresponde al número de niños que repitieron en cada año.

El cuadro 4 presenta los riesgos acumulados al tercer año ($1-S_{t=3}$), condicionados a las variables de interés, incluida la valoración global del desempeño en EIT. Se incluyen, además, los estadísticos descriptivos básicos para cada variable (media o proporción).

El análisis descriptivo muestra, en primer lugar, que los riesgos de repetición están altamente asociados al nivel de habilidades tempranas. Son de $q=0.569$ para los estudiantes valorados como rojos por EIT a los cinco años, $q=0.329$ para los amarillos y $q=0.081$ para los verdes. Ninguna de las restantes variables consideradas muestra una asociación tan fuerte con la repetición, aunque muchas discriminan los riesgos de interés de manera importante y en el sentido previsto. Por ejemplo, las probabilidades de repetir durante los primeros tres años en primaria se incrementan entre los varones, entre los niños de menor edad de su generación, entre los alumnos con mayor ausentismo a clases en inicial, entre quienes asisten a escuelas de menor nivel socioeconómico, entre los que provienen de familias con menos educación y entre los que registran indicadores perinatales considerados críticos.

■ Cuadro 4. Indicadores y estadísticos descriptivos

| Variable | Descripción | Proporción o media | 1-5 _{t=3} |
|---|---|--------------------|--------------------|
| EVALUACIÓN INFANTIL (EIT) ⁽¹⁾ | | | |
| Resumen = Rojo | Indicador habilidad global EIT | 0.094 | 0.559 |
| Resumen = Amarillo | | 0.144 | 0.329 |
| Resumen = Verde | | 0.763 | 0.082 |
| VARIABLES CONTEXTUALES | | | |
| CSE de la escuela: Quintil 1 ⁽²⁾ | Proporción de la muestra en escuelas de quintil 1 (NSE) | 0.183 | 0.247 |
| CSE de la escuela: Quintil 5 ⁽²⁾ | Proporción en quintil 5 (NSE) | 0.232 | 0.094 |
| AFAM_PE1 ⁽³⁾ | Hogar beneficiario de AFAM_PE | 0.557 | 0.206 |
| TUS1 ⁽³⁾ | Hogar beneficiario de TUS | 0.254 | 0.268 |
| m_niveled2 ⁽⁴⁾ | Educación materna: Primaria completa o menos | 0.222 | 0.261 |
| m_niveled6 ⁽⁴⁾ | Educación materna: Terciaria | 0.065 | 0.032 |
| VARIABLES PERINATALES | | | |
| Edadma ⁽⁵⁾ | Edad de la madre al parto: 19 o menos | 0.198 | 0.188 |
| pesor2 ⁽⁵⁾ | Bajo peso al nacer | 0.081 | 0.256 |
| numemb ⁽⁵⁾ | Nro de embarazo: 4 o más | 0.111 | 0.267 |
| gesta ⁽⁵⁾ | Semanas de gestación: 36 o menos | 0.083 | 0.240 |
| convpad ⁽⁵⁾ | Padre no convive al momento del parto | 0.339 | 0.217 |
| controles ⁽⁵⁾ | Nro de controles en el embarazo: 5 o menos | 0.169 | 0.273 |
| captacion ⁽⁵⁾ | Captación del embarazo: semana 15 o más | 0.321 | 0.212 |
| VARIABLES SOCIODEMOGRÁFICAS | | | |
| Edad (primer cuartil) ⁽⁴⁾ | 25% de menor edad de la generación | 0.251 | 0.200 |
| Edad (4to cuartil) ⁽⁴⁾ | 25% de mayor edad de la generación | 0.248 | 0.131 |
| Niñas ⁽⁴⁾ | Proporción de niñas | 0.495 | 0.127 |
| Varones ⁽⁴⁾ | Proporción de varones | 0.505 | 0.196 |
| INASISTENCIAS EN EDUCACIÓN INICIAL | | | |
| faltas_2016 (primer cuartil) ⁽⁴⁾ | 25% con menor cantidad de faltas en NI 5 (2016) | 0.272 | 0.094 |
| faltas_2016 (4to cuartil) ⁽⁴⁾ | 25% con mayor cantidad de faltas en NI 5 (2016) | 0.248 | 0.269 |

Fuente: elaboración propia con base en datos de ⁽¹⁾ DIEE-ANEP; ⁽²⁾ CEIP-ANEP; ⁽³⁾ SIAS-MIDES; ⁽⁴⁾ GURI-CEIP; ⁽⁵⁾ CNV-MSP.

Regresión logística

El primer modelo de riesgo, correspondiente a la especificación logística, arroja tres resultados principales. El modelo muestra, en primer lugar, que la fuerte asociación reportada con base en el análisis descriptivo entre las habilidades captadas por EIT en el nivel 5 de inicial y los riesgos de repetición en los tres años posteriores se mantiene luego de controlar por el resto de las variables. *Ceteris paribus*, la probabilidad de que un niño se rezague en su trayectoria escolar en ese tiempo es 15.6 puntos porcentuales más alta para los alumnos valorados como amarillos en EIT (dificultades moderadas) y 25.1 puntos mayor para los rojos (dificultades severas), en comparación con los verdes (sin dificultades).

En segundo lugar, el modelo muestra asociaciones significativas y en el sentido esperado para varios de los restantes indicadores considerados, vinculados a las dimensiones contextual (hogar TUS, contexto de la escuela, educación materna), sociosanitaria (bajo peso al nacer, multiparidad, la ausencia en el hogar del padre al momento del nacimiento y controles durante el embarazo), sociodemográfica (sexo y edad en meses), así como para el ausentismo a clases en el último año del nivel inicial. Los efectos marginales de estas variables sobre los riesgos de repetición son, de todos modos, sustantivamente más bajos en comparación con los reportados para el indicador global de habilidades tempranas relevados por EIT.

Nótese que las estimaciones obtenidas mediante el modelo de regresión corresponden exclusivamente a los efectos directos de cada variable sobre los riesgos de repetición. En particular, al controlar por el desempeño en EIT, esta especificación invisibiliza posibles efectos indirectos de los otros predictores del tipo $x \rightarrow \text{EIT} \rightarrow Y$, que podrían estar impactando en el resultado de interés de forma mediada, es decir, a través de sus efectos previos sobre el desarrollo de las habilidades tempranas.⁶ La especificación mediante redes bayesianas, que presentamos en el cuadro 5, permite una representación más adecuada en este sentido.

⁶ En particular, tal como hemos mostrado en un trabajo anterior (Borba *et al.*, 2018), las habilidades captadas por EIT al término de la educación inicial están estrechamente asociadas a las circunstancias sociosanitarias al nacimiento, a las condiciones socioeconómicas “contemporáneas” a la evaluación EIT y a atributos sociodemográficos como el sexo y la edad.

■ Cuadro 5. Modelo de regresión logística. Efectos marginales

| | Marginal effects | ee | sig |
|-----------------------------|------------------|-------|------|
| VARIABLES CONTEXTUALES | | | |
| quintil1 | -0.032 | 0.007 | (**) |
| AFAM_PE1 | 0.011 | 0.008 | |
| TUS1 | 0.027 | 0.008 | (**) |
| m_niveled2 | -0.029 | 0.016 | |
| m_niveled4 | -0.067 | 0.016 | (**) |
| m_niveled5 | -0.099 | 0.020 | (**) |
| m_niveled6 | -0.165 | 0.027 | (**) |
| VARIABLES PERINATALES | | | |
| edadma191 | 0.015 | 0.009 | |
| pesor21 | 0.032 | 0.013 | (**) |
| numemban1 | 0.042 | 0.008 | (**) |
| gesta1 | 0.001 | 0.013 | |
| convpad1 | 0.046 | 0.006 | (**) |
| controles1 | 0.029 | 0.008 | (**) |
| captacion1 | 0.006 | 0.007 | |
| VARIABLES SOCIODEMOGRÁFICAS | | | |
| qedad1 | -0.021 | 0.007 | (**) |
| na1 | -0.033 | 0.006 | (**) |
| AUSENTISMO ESCOLAR | | | |
| faltas_20161 | 0.057 | 0.007 | (**) |
| EVALUACIÓN INFANTIL | | | |
| Resumen = Rojo | 0.251 | 0.008 | (**) |
| Resumen = Amarillo | 0.156 | 0.007 | (**) |

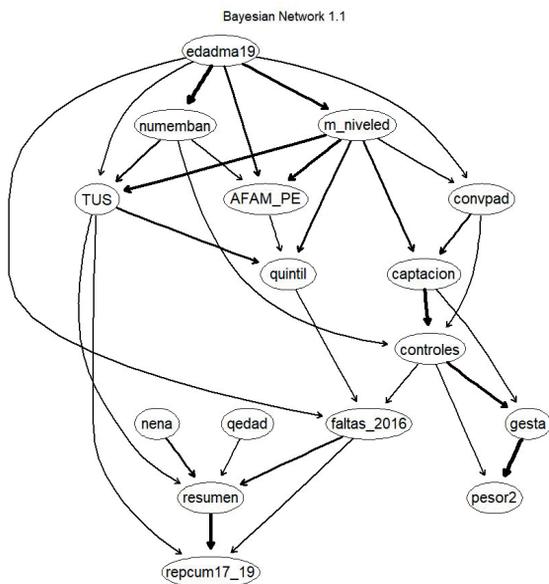
Fuente: elaboración propia con base en Panel-EIT; registros del CNV-MSP, SIIAS-MIDES y ANEP. (**) Efectos significativos al 95%.

MODELO DE REDES BAYESIANAS (BN)

La figura 1 presenta el GAD correspondiente al modelo de redes bayesianas (BN) estimado. Subrayamos tres resultados principales de este segundo modelo predictivo. Primero, al igual que en el modelo de regresión, las habilidades globales captadas por EIT vuelven a aparecer como el mejor predictor de los riesgos de repetición posterior.

Esto se refleja, gráficamente, en el grosor de la flecha que conecta este indicador con la variable de resultado y aparece cuantificado en el cuadro 6, que reporta la magnitud de las asociaciones estimadas.

■ Figura 1. Modelo redes bayesianas con indicador resumen de EIT



Fuente: elaboración propia con base en el Panel EIT 2016-22; registros del CNV-MSP, SIAS-MIDES y ANEP

Segundo, además de las habilidades EIT, el GAD identifica sólo otras dos variables padres (es decir, con efectos directos) de la repetición: la participación del hogar en el programa TUS y el ausentismo a clases durante el año previo a la transición a primaria. El modelo sugiere que, además de sus efectos directos sobre los riesgos de repetición, ambas variables inciden indirectamente a través de su impacto previo en el nivel de habilidad global captado por EIT.

- i. Educación de la madre → hogar beneficiario de programas sociales → nivel socioeconómico contemporáneo de la familia → ausentismo a clases en educación inicial → habilidades tempranas → riesgo de reptición;

- ii. *Educación de la madre* → semana de captación del embarazo → número de controles durante el embarazo → ausentismo a clases en educación inicial → habilidades tempranas → riesgo de reptición;
- iii. *Edad* → habilidades tempranas → riesgo de reptición;
- iv. *Sexo* → Habilidades tempranas → riesgo de reptición ;

Esta estructura de mediaciones no implica que la capacidad predictiva que tienen las habilidades tempranas sobre la repetición esté reflejando de modo exclusivo el efecto indirecto de las variables antecesoras. Aunque escapa a las posibilidades del presente análisis, una hipótesis en términos causales desde las habilidades tempranas hacia los logros educativos posteriores es, al menos, verosímil, en especial si se considera que la activación de estas habilidades constituye un requerimiento necesario para la adaptación a la escuela o para el aprendizaje de habilidades más complejas, como la escritura.

El cuadro 6 reporta los riesgos estimados por el ajuste a partir de la ecuación (4) definida en la sección anterior, con base en los tres predictores identificados. Como puede observarse, los riesgos se incrementan en el sentido esperado en función de cualquiera de las tres variables, independientemente del valor de las dos restantes. Estos incrementos son en particular pronunciados en relación con el indicador de habilidad global de EIT, con brechas de entre 42 y 50 puntos entre los niños valorados como rojos y como verdes, según su situación en los otros dos indicadores.

■ Cuadro 6. Riesgos de repetición estimados por BN

| | | FALTAS en Nivel 5 (2016) | |
|----------------|----------|--------------------------|-------|
| | | <50 | 50+ |
| Hogar TUS = No | Verde | 0.053 | 0.138 |
| | Amarillo | 0.262 | 0.399 |
| | Rojo | 0.517 | 0.559 |
| Hogar TUS = Sí | Verde | 0.137 | 0.200 |
| | Amarillo | 0.385 | 0.405 |
| | Rojo | 0.565 | 0.696 |

Fuente: elaboración propia con base en el Panel EIT 2016-22; registros del CNV-MSP, SIAS-MIDES y ANEP

Árbol de clasificación

La figura 2 presenta el resultado del tercer modelo de riesgo (árbol de clasificación). El árbol comienza con un único nodo, que corresponde a 100% de las observaciones. En cada partición sucesiva se reporta: el riesgo q estimado y el porcentaje de casos clasificados en relación con el total de la muestra para el grupo correspondiente. Los árboles ofrecen una forma gráfica para visualizar diferentes configuraciones de riesgo, lo que constituye un enfoque complementario a la noción de “efectos netos”, de carácter aditivo, propia de las técnicas de regresión.

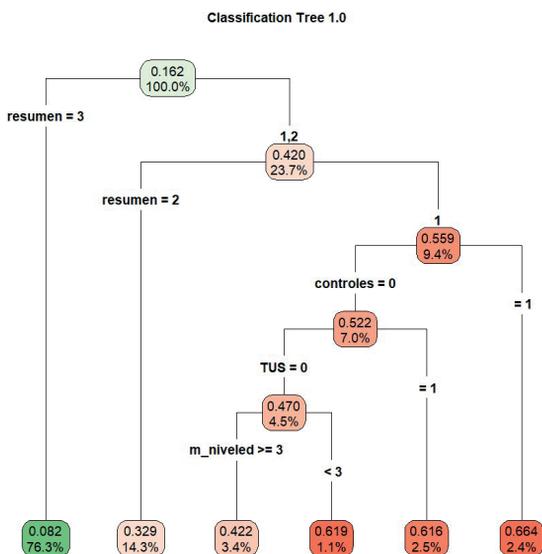
Como se desprende de la figura 2, el indicador resumen de las habilidades captadas por EIT estructura de manera predominante la solución del árbol de clasificación. De hecho, las primeras dos particiones corresponden a la distinción, primero, entre los alumnos valorados por EIT como verdes (76.3% de la cohorte, con un riesgo estimado $q=0.082$) y el resto ($q=0.420$; 27.7%) y, para este último grupo, en un segundo paso, entre los valorados como amarillos (14.3%; $q=0.329$) y los rojos (9.4%; $q=0.559$).

Los alumnos categorizados como verdes y amarillos dan lugar a las dos primeras hojas del árbol (en la base de la figura, de izquierda a derecha), sin necesidad de considerar ninguna variable adicional. En el grupo de los rojos, en tanto, la solución implica una sucesión de particiones adicionales que involucran, en orden, los siguientes indicadores: el número de controles de la madre durante el embarazo, el ser o no beneficiario del programa TUS y el nivel educativo de la madre del alumno. Estas particiones dan lugar a cuatro nodos u hojas finales.

A efectos específicos de la clasificación de los alumnos dentro del grupo de riesgo, y dado el umbral definido ($\pi=0.20$), la única distinción relevante que surge de este último modelo es entre el primer grupo (alumnos valorados como verdes por EIT) y todos los restantes (amarillos y rojos), dado que, para estos últimos, $q > \pi$, independientemente de la consideración de las restantes variables.⁷

⁷Dado que el desempeño global en EIT termina siendo el único criterio relevante para la predicción, las brechas en los riesgos de repetición entre los niños valorados como rojos, amarillos y verdes estimadas por el árbol de clasificación coinciden con las probabilidades de repetición respectivas, reportadas en el cuadro 4.

■ Figura 2. Árbol de clasificación



Fuente: elaboración propia con base en el Panel EIT 2016-22; registros del CNV-MSP, SIIAS-MIDES y ANEP

Comparación del ajuste global de los modelos

La primera columna del cuadro 7 reporta los indicadores de ajuste global para los tres modelos de riesgo analizados en la sección anterior. En la columna (2) se estiman estos mismos indicadores, pero omitiendo de los modelos correspondientes, a efectos comparativos, al indicador de habilidad global en EIT. Al final de la tabla, se calculan los mismos indicadores de ajuste para una predicción bivariada realizada en exclusivo con base en el desempeño en EIT.⁸

⁸ Para este ejercicio se considera en riesgo a todos los alumnos valorados como rojos o amarillos por EIT.

■ Cuadro 7. Indicadores de ajuste de los modelos de riesgo escolar (regresión logística; árbol de clasificación y redes bayesianas). Resultado de interés: al menos un evento de repetición entre primero y tercero de primaria^(a)

| | (1) Modelos con EIT | (2) Sin considerar el desempeño global en EIT |
|---|---------------------|---|
| i) Regresión logística | | |
| Precisión | 0.80 | 0.72 |
| Kappa | 0.39 | 0.23 |
| Sensibilidad | 0.64 | 0.56 |
| Especificidad | 0.83 | 0.75 |
| AUC ^(b) | 0.81 | 0.73 |
| ii) Redes bayesianas ^(c) | | |
| Precisión | 0.80 | 0.70 |
| Kappa | 0.38 | 0.20 |
| Sensibilidad | 0.62 | 0.55 |
| Especificidad | 0.83 | 0.73 |
| iii) Árbol de clasificación | | |
| Precisión | 0.80 | 0.71 |
| Kappa | 0.38 | 0.16 |
| Sensibilidad | 0.62 | 0.42 |
| Especificidad | 0.83 | 0.77 |
| iv) Predicción bivariada con base en EIT ^(d) | | |
| Precisión | 0.80 | -- |
| Kappa | 0.38 | -- |
| Sensibilidad | 0.62 | -- |
| Especificidad | 0.84 | -- |

Fuente: elaboración propia con base en el Panel-EIT; registros del CNV-MSP, SIAS-MIDES y ANEP. ^(a) El umbral para definir un caso como positivo, en todos los modelos, se definió por $\hat{p}(x) \geq 0.20$. ^(b) El indicador AUC se reporta únicamente para el método de regresión. ^(c) Método de Hill-climbing (BIC score). ^(d) Clasifica como en riesgo a los alumnos valorados como rojos o amarillos por EIT.

Destacamos tres resultados. Primero, los modelos logran clasificar correctamente a 80% del total de los alumnos (*precisión*), lo que supone una mejora sustantiva respecto a una predicción basada de manera exclusiva en el azar (*kappa*). Además, identifican a entre 62% y 64% de los niños que efectivamente terminaron rezagándose entre primero y tercero (*sensibilidad*) e incluyen “erróneamente” dentro del grupo de riesgo a menos de 20% (1- *especificidad*). El desempeño predictivo de los tres modelos es prácticamente idéntico; es decir, los resultados son robustos al método (columna 1).

Segundo, la inclusión del desempeño en EIT captado a los cinco años mejora de manera sustantiva la capacidad predictiva de los modelos de riesgo en todos los indicadores de ajuste e independientemente del método de clasificación (columna 1 *vs.* columna 2). Este resultado es consistente con los análisis presentados en las secciones anteriores y refuerza la hipótesis del alto poder predictivo de las habilidades tempranas respecto a los riesgos de repetición en primaria.

Tercero, aunque incluyen una batería muy importante de indicadores sociodemográficos, sociosanitarios, contextuales y educativos, fuertemente asociados con la repetición en la escuela, los modelos de riesgo estimados no suponen una mejora sustantiva respecto a la predicción que surge de una matriz de confusión basada de manera exclusiva en el desempeño global en EIT. Esto ilustra que, a efectos estrictamente predictivos, el resto de las variables consideradas en los tres modelos multivariados de riesgo aporta información redundante, una vez que se conoce el nivel de habilidades de los niños al término de la educación inicial.

DISCUSIÓN E IMPLICANCIAS

Discusión

Los análisis presentados en este artículo se focalizaron en la valoración de criterios de clasificación de los niños de cinco años según su riesgo escolar futuro; en particular, procuramos valorar la capacidad predictiva que por sí sola, y en conjunto con una batería amplia de indicadores, tienen las habilidades globales demostradas a esa edad, respecto a los riesgos de experimentar un evento de repetición durante los primeros tres años en la escuela graduada.

Tal como esperábamos, los riesgos de repetición varían de modo sustantivo en función de las habilidades tempranas vinculadas a la PPE. De hecho, nuestros resultados indican que el indicador global de EIT logra, por sí solo, una clasificación de los niños en riesgo de rezagarse más robusta que la que deriva de la consideración individual o conjunta de buena parte de las variables que la literatura especializada identifica como factores asociados al logro educativo, entre los que evaluamos: factores de tipo contextual, como el nivel socioe-

conómico de la escuela y del hogar; condiciones sociosanitarias del niño y de su familia al momento del nacimiento; rasgos sociodemográficos exógenos, como el sexo y la edad, y el nivel de ausentismo a la escuela durante el último año en la educación inicial.

Los modelos predictivos que presentamos en este artículo llegan a resultados similares y muestran la preponderancia que tiene la medida global de habilidades vinculadas a la PPE para la predicción de la repetición escolar y, en consecuencia, su alto potencial como criterio de clasificación temprana de niños en situación de riesgo educativo. Los tres ajustes considerados, basados en modelos de regresión logística, de redes bayesianas y de árboles de clasificación, logran niveles prácticamente idénticos de precisión, sensibilidad y especificidad: clasifican de manera correcta a 80% de los alumnos y, lo que es más importante, aciertan a identificar a entre 62-64% de los niños que, en efecto, terminan rezagándose entre primer y tercer año de primaria (sensibilidad). Tal como se mostró, a pesar de que hacen un uso intensivo de información sobre una batería amplia de factores asociados, ninguno de estos modelos supone una mejora sustantiva, en términos predictivos, respecto a una clasificación basada de modo exclusivo en el desempeño global de EIT.

Este último resultado no debe llevar a subestimar los impactos que sobre la trayectoria escolar tienen factores como el contexto socioeconómico y cultural de los niños, las condiciones sociosanitarias al nacimiento y durante la primera infancia, la edad, el sexo o el nivel de ausentismo a clases en la educación inicial. De hecho, tal como surge en especial de los modelos de regresión y de redes bayesianas, en términos sustantivos, varios de estos factores se encuentran en la base de las diferencias observadas en el nivel de habilidades desarrolladas hasta los cinco años, lo que sugiere que sus impactos sobre los resultados escolares posteriores no son insignificantes, sino que actúan en buena medida en forma indirecta, es decir, mediados por dichas habilidades.

Implicancias

Nuestros resultados muestran que es posible identificar con anticipación y con niveles altos de precisión, sensibilidad y especificidad

a aquellos niños que, con alta probabilidad, enfrentarán dificultades para progresar en la escuela en los ritmos normativamente previstos. Además, sugieren que la identificación temprana de estos niños puede hacerse de manera escalable al conjunto de la matrícula y mediante protocolos de valoración relativamente sencillos, como los desarrollados por EIT (o, actualmente, por INDI) en Uruguay, aplicados por los propios docentes en contextos cotidianos de clase.

Más aún, la evidencia indica que la información que arroja la aplicación de un instrumento como EIT al término de la educación inicial es suficientemente potente, por sí misma, para alcanzar una alta potencia predictiva, incluso sin considerar otras variables de más compleja medición o que se asocian a los resultados educativos de manera más indirecta. En otros términos, la valoración de las habilidades tempranas aporta, por sí misma, el mejor criterio de clasificación temprana de niños en riesgo escolar.

Más allá del poder predictivo, desde el punto de vista de la política educativa, basar un sistema de alertas tempranas en indicadores vinculados al desarrollo infantil y a las habilidades asociadas a la PPE resulta una alternativa sumamente interesante, al menos por tres razones complementarias. En primer lugar, un sistema de este tipo se basaría en un criterio de clasificación específico: a diferencia de otros factores que típicamente se utilizan como predictores del riesgo escolar —por ejemplo, el nivel socioeconómico de la familia o el entorno socioeconómico de la escuela—, los déficits asociados al desarrollo de las habilidades infantiles están directamente vinculados a la PPE y, por tanto, aportan pautas para la intervención pedagógica directa sobre aquellas dificultades que generan el riesgo de interés; es decir, constituyen un criterio estrechamente asociado a la posibilidad de desarrollar intervenciones por niveles de respuesta. En segundo lugar, se basa en criterios fácilmente comunicables, tanto a la comunidad de docentes y técnicos que se desempeñan en las escuelas y al sistema educativo en general, como a las propias familias involucradas. Tercero, un sistema basado en la valoración de las habilidades tempranas permitiría complementar los criterios predominantes de focalización de la política educativa —que en la mayoría de los países, incluido Uruguay, han tomado al centro educativo como unidad básica de la intervención—, mediante estrategias individualizadas, basadas en la clasificación temprana de niños según

su riesgo educativo y en la definición y desarrollo de intervenciones de carácter personalizado, enfocadas en aquellos alumnos que así lo requieran, independientemente del tipo o contexto de la escuela a la que asisten.

REFERENCIAS

- Allotey, J., J. Zamora, F., Cheong-See, M., Kalidindi, D., Arroyo-Manzano, E., Asztalos K., y Khan, S. (2017). Cognitive, motor, behavioural and academic performances of children born preterm: a meta-analysis and systematic review involving 64.061 children. *BJOG: An International Journal of Obstetrics & Gynaecology*, 125(1), 16-25. <https://www.doi.org/10.1111/1471-0528.14832>
- Almond, D., y Currie, J. (2010). *Human capital development before age five*. Working Paper 15827. Cambridge, Massachusetts: NBER.
- Amarante, V., y Labat, J.P. (2018). *Las políticas públicas dirigidas a la infancia: aportes desde el Uruguay (LC/TS.2018/68-LC/MVD/TS.2018/3)*. Santiago de Chile: Comisión Económica para América Latina y el Caribe (CEPAL).
- Arrivillaga, C., Cuevasanta, D., Liz, M., Moreira, K., Schiappacasse, P., y Vásquez, A. (2016). Preparación para la escolarización: una revisión sistemática de estudios longitudinales. *PSIENCIA. Revista Latinoamericana de Ciencia Psicológica*, 8(1), 1-12.
- Bhattacharjea, S. (2019). Are Schools “Ready for Children”? Assumptions and Ground Realities. En V. Kaul y S. Bhattacharjea (eds.), *Early Childhood Education and School Readiness in India* (pp. 195-213). Singapur: Springer. https://doi.org/10.1007/978-981-13-7006-9_10
- Blair, C. (2001). The early identification of risk for grade retention among African American children at risk for school difficulty. *Applied Developmental Science*, 5(1), 37-50.
- Borba, E., Cardozo, S., Castelao, G., Peri, A., Salsamendi, G., y Silveira, A. (2018). *Desarrollo infantil y riesgo de repetición en primero*. Montevideo: DIEE-ANEP.
- Cabella, W., De Rosa, M., Failache, E., Fitermann, P., Katzkowicz, N., y Medina, M. (2015). *Salud, nutrición y desarrollo en la primera infancia en Uruguay: primeros resultados de la ENDIS*. Montevideo: INE, UDELAR, OPP, MS, MIDES.

- Christopolou, M., y Pampaka, M. (2015). Assessing school readiness. Cyprus, *Case Studies Journal*, 4(10).
- Clark, C., Pritchard, V., y Woodward, L. (2010). Preschool executive functioning abilities predict early mathematics achievement. *Developmental Psychology* 46(5), 1176-1191. <https://www.doi.org/10.1037/a0019672>
- Davoudzadeh, P., McTernan, M. L., y Grimm, K. J. (2015). Early school readiness predictors of grade retention from kindergarten through eighth grade: A multilevel discrete-time survival analysis approach. *Early Childhood Research Quarterly*, 32(3), 183-192.
- Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., y Japel, C. (2007). School readiness and later achievement. *Developmental Psychology*, 43(6), 1428-1446.
- Failache, E., y Katzkowicz N. (2018). Desarrollo infantil: Una aproximación a sus determinantes. *Concurso de proyectos de investigación: "Primera infancia: análisis comparado de la primera y segunda ola de la Encuesta de Nutrición, Desarrollo Infantil y Salud (ENDIS)"*. Ministerio de Desarrollo Social, Montevideo, Uruguay.
- Failache, E., Salas G., y Vigorito A. (2015). *Trayectorias educativas de los adolescentes en Uruguay. Un estudio en base a datos de panel*. Serie Documentos de Trabajo, DT 06/2015. Uruguay: Instituto de Economía, FCEA-UDELAR.
- Gledhill, J., Ford, T., y Goodman, R. (2002). Does season of birth matter?: The relationship between age within the school year (season of birth) and educational difficulties among a representative general population sample of children and adolescents (aged 5-15) in Great Britain. *Research in Education*, 68, 41-47.
- Hair, E., Halle, T., Terry-Humen, E., Lavelle, B., y Calkins, J. (2006). Children's school readiness in the ECLS-K: Predictions to academic, health, and social outcomes in first grade. *Early Childhood Research Quarterly*, 21(4), 431-454.
- Jung, E. (2016). The Development of Reading Skills in Kindergarten Influence of Parental Beliefs About School Readiness, Family Activities, and Children's Attitudes to School, *International Journal of Early Childhood*, 48(1), 61-78. <https://www.doi.org/10.1007/s13158-016-0156-2>

- La Paro, K. M., y Pianta, R. C. (2000). Predicting children's competence in the early school years: A meta-analytic review. *Review of Educational Research*, 70(4), 443-484.
- Lee, V. E., y Burkam, D. T. (2002). *Inequality at the Starting Gate: Social Background Differences in Achievement as Children Begin School*. Seattle: Economic Policy Institute.
- Linan-Thompson, S. (2014). *Screening of school readiness skills: A review of the literature*. Washington, D. C.: USAID y Global: JBS International. https://pdf.usaid.gov/pdf_docs/PA00JW1P.pdf
- Liz, M. (2018). *Autorregulación y preparación para la escolarización. El funcionamiento ejecutivo en educación inicial predice el involucramiento escolar en primer año* (Tesis de Maestría). Universidad de la República, Uruguay.
- López, A., y Salsamendi, G. (2018). La universalización de la Evaluación Infantil Temprana: una apuesta a la equidad. *III Congreso Latinoamericano de Medición y Evaluación Educacional (COLMEE)*, Montevideo.
- López, A., y Willms, D. (2020). A National Evaluation of Kindergarten Outcomes: Findings from Uruguay. En J. Hall, A. Lindorff y P. Sammons (eds.), *International Perspectives on Educational Effectiveness Research* (pp. 361-381). Nueva York: Springer.
- Luu, T. M., Mian, M.O.R, y Nuyt, A. M. (2017). Long-Term Impact of Preterm Birth. *Clinics in Perinatology*, 44(2), 305-314.
- Mara, S., Alesina, L., Cabrio, S., Erramouspe, R., Pazos, L., e Ibañez, W. (2000). *Estudio de Evaluación de impacto de la Educación Inicial en el Uruguay*. Montevideo: ANEP.
- National Early Literacy Panel (2008). *Developing early literacy: Report of the National Early Literacy Panel*. Washington, DC: National Institute for Literacy. <http://lincs.ed.gov/publications/pdf/NELPReport09.pdf>
- Pan, Q., Trang, K. T., Love, H. R., y Templin, J. (2019). School Readiness Profiles and Growth in Academic Achievement. *Frontiers in Education*, 4(127). <https://www.doi.org/10.3389/feduc.2019.00127>
- Raver, C., S. Jones, S., Li-Grining, C., Zhai, F., Bub, K., y Pressler, E. (2011). The Chicago School Readiness Project's Impact on Low-Income Preschoolers' Preacademic Skills: Self-regulation as a Mediating Mechanism. *Child Development*, 82, 362-378. <https://www.doi.org/10.1111/j.1467-8624.2010.01561.x>

- Rebello, P. (2012). *School Readiness. A Conceptual Framework*. Nueva York: UNICEF.
- Rimm-Kaufman, S., y Sandilos, L. (2017). School Transition and School Readiness: An Outcome of Early Childhood Development. En R. Tremblay, M. Boivin y R. Peters (eds.), *Encyclopedia on Early Childhood Development*. <http://www.child-encyclopedia.com/school-readiness/according-experts/school-transition-and-school-readiness-outcome-early-childhood>
- Scarborough, H. S. (2001). Connecting early language and literacy to later reading (dis)abilities: Evidence, theory and practice. En S. B. Neuman y D. K. Dickinson (eds.), *Handbook of early literacy research* (pp. 97–110). Nueva York: Guilford Press.
- Sloat, E. A., Beswick, J. F., y Willms, J.D. (2007). Using early literacy monitoring to prevent reading failure. *Phi Delta Kappan*, 88(7), 523-529.
- UNESCO (2018). *Learning Divides: Using Data to Inform Educational Policy*. Information Paper No. 54. Montreal: UNESCO Institute for Statistics.
- UNICEF (2012). *School Readiness: A Conceptual Framework*. Nueva York: Fondo de las Naciones Unidas para la Infancia.
- Vásquez, A., y Moreira, K. (2016). *Preparación para la escolarización. Dimensiones y Medición*. Uruguay: Facultad de Psicología, Universidad de la República.
- Willms, D., Laurie, R., Tunison, S., y Haley, A. (2000). *Using Early Years Evaluation in a Response-to-Intervention (RTI) Model*. Estados Unidos: Research, Evaluation, and Assessment in Schools.
- Wilson, S. J. (2014). *School Readiness and Later Achievement: Results from a Meta-analysis of Longitudinal Studies*. SREE Spring Conference, Washington, D. C.