

Alertas tempranas para prevenir el abandono escolar: el caso de la provincia de Mendoza



Alertas tempranas para prevenir el abandono escolar: el caso de la Provincia de Mendoza

© CAF 2023

Publicación desarrollada en el marco de la asistencia técnica de CAF a la Provincia de Mendoza de Apoyo al Desarrollo de un “Modelo Predictivo para un Sistema de Alerta Temprana de prevención del abandono escolar”.

Coordinación de la asistencia técnica y publicación:

Por CAF: Cecilia Llambí y Lucila Berniell

Por la Dirección General de Escuelas de Mendoza (DGE):
Romina Paola Durán

Autores:

Por CAF: Lucila Berniell, Agustina Hatrick y Cecilia Llambí

Por DGE Mendoza: Romina Paola Durán, Margarita Olivera,
Leopoldo Javier Ontivero y Paula Ortega Grebenc

Diseño y maquetación: Humaga | www.humaga.com.ar

Agradecimientos: Al Laboratorio de Inteligencia Artificial Aplicada (LIAA) de la Universidad de Buenos Aires (UBA), en particular a Juan Kamienkowski, Pablo Riera y Gustavo Juantorena.

Las ideas y opiniones expresadas en esta obra son las de los autores y no reflejan necesariamente el punto de vista de CAF ni comprometen a la organización.

Disponible en la biblioteca digital de CAF scioteca.caf.com con acceso abierto bajo la licencia Reconocimiento-NoComercial-SinObraDerivada 4.0 Internacional (CC-BY-NC-ND 4.0)



introducción

Antes de la pandemia del COVID-19, América Latina ya enfrentaba un grave déficit de aprendizaje, evidenciado en los resultados de pruebas internacionales. Por ejemplo, en las Pruebas PISA 2018 y PISA-D, solo el 30% de los y las estudiantes de 15 años en América Latina tenía competencias mínimas en matemática. Las oportunidades de aprendizaje no son equitativas, ya que existe una fuerte relación entre el nivel socioeconómico y el rendimiento académico. En América Latina, la brecha en el rendimiento en matemática entre los hogares más y menos privilegiados asciende a casi 40 puntos porcentuales, en contraste con los 28 puntos porcentuales en los países de la OCDE. Solo el 14% de los y las jóvenes de hogares más desfavorecidos económicamente alcanza competencias mínimas en matemática (Berniell et al, 2021).

Esta disparidad socioeconómica no solo se refleja en el rendimiento de los estudiantes, sino también en el acceso y la permanencia en la escuela. Si bien las brechas en la tasa de matriculación en la educación primaria y secundaria entre los niños de hogares más ricos y más pobres se redujeron en las últimas tres décadas, estas brechas aún persisten y se mantienen altas en varios países de la región. A su vez, la evidencia sugiere que la brecha socioeconómica en el aprendizaje y en la permanencia en el sistema educativo pudo haberse ampliado luego de la pandemia, proyectándose incrementos en la desvinculación y el abandono educativo (UNESCO, CEPAL y UNICEF, 2022).

Muchos de los y las estudiantes que inician la educación secundaria no la completan y en consecuencia nunca se gradúan. Antes de la pandemia se estimaba que alrededor de 92,7% de los niños y las niñas de la región concluía la primaria, pero solo el 63,7% lograba graduarse de la secundaria. Y mientras que el 84,6% de los y las jóvenes del quintil de ingresos más alto logra graduarse de la educación secundaria, solamente el 44,1% de los y las provenientes del quintil de ingresos más bajos lo logra (UNESCO, CEPAL y UNICEF, 2022). Frecuentemente, los y las jóvenes no se sienten apoyados en sus

escuelas, no aprenden lo suficiente y no visualizan cómo la educación que reciben los prepara para el futuro. A su vez, el embarazo adolescente y las dificultades económicas continúan siendo obstáculos importantes para que los y las jóvenes logren transitar con éxito la educación obligatoria (Berniell et. al, 2016). Estos obstáculos afectan la trayectoria escolar y ello tiene un impacto en el resto de sus vidas: no culminar la escuela secundaria suele asociarse a tener menores ingresos, menores probabilidades de conseguir y mantener un trabajo y mayor propensión a incurrir en conductas riesgosas.

En ese contexto, varios países de la región comenzaron a desarrollar Sistemas de Alerta Temprana (SAT) para prevenir el abandono escolar. Estos sistemas se centran en la prevención oportuna, lo que implica identificar con la suficiente antelación y con la mayor precisión posible a los y las estudiantes que están en riesgo de interrumpir su trayectoria educativa, con el propósito de acompañar sus trayectorias con acciones que permitan fortalecerlas.

Este documento presenta la iniciativa del diseño e implementación de un Sistema de Alerta Temprana en la provincia de Mendoza, Argentina, desarrollado sobre la base de su sistema de gestión educativa, Gestión Educativa Mendoza (GEM). En la sección I se describe el contexto y las políticas de la provincia de Mendoza que tienen como herramienta fundamental el GEM y el uso de la información para la toma de decisiones, así como la problemática del abandono escolar. En la sección II se presenta conceptualmente qué es un SAT y cuáles son sus principales elementos. A partir de ello, en la sección III se presenta el SAT elaborado a partir de la información del GEM. Finalmente, en la sección IV se concluye sobre las oportunidades y desafíos para la implementación de este tipo de políticas.

Gestión basada en evidencia: el recorrido de la provincia de Mendoza

Antesala del SAT: el uso de datos para la gestión en la actualidad

Desde el gobierno escolar de la Provincia de Mendoza se ha establecido una estrategia fundamental para alcanzar sus objetivos: fortalecer el seguimiento de las trayectorias de las y los estudiantes, para poder identificar con mayor certeza y precisión las necesidades de cada una/o y su progreso en el sistema educativo.

La herramienta provincial que permite este seguimiento nominalizado es el Sistema de Gestión Educativa de Mendoza (GEM). Este sistema, garantiza a la gestión educativa datos confiables y actualizados de cada estudiante, lo que permite realizar un seguimiento preciso y un monitoreo efectivo de todas las trayectorias escolares.

Además de contribuir al cumplimiento de cada eje de la política educativa, la información proporcionada por el GEM funciona como un gran potencial que brinda la posibilidad de diseñar estrategias para abordar la problemática del abandono y la exclusión escolar.

ESQUEMA 1. EJES DE LA POLÍTICA EDUCATIVA DE LA DIRECCIÓN GENERAL DE ESCUELAS (DGE) 2020 – 2023



Sistema de Gestión Educativa de Mendoza - GEM

Desde marzo de 2017, Mendoza cuenta con un sistema de gestión escolar digital llamado GEM (Gestión Educativa Mendoza). Este sistema fue creado por la misma Dirección General de Escuelas, con la finalidad de ofrecer a todas las instituciones de gestión estatal y privada una herramienta tecnológica actualizada que permita facilitar, agilizar y organizar la gestión administrativa-pedagógica escolar de manera dinámica, colaborando al proyecto de modernización del Estado de la Provincia de Mendoza y aportando transparencia a la gestión.

En la actualidad, el GEM mantiene un objetivo primordial: unificar los distintos sistemas y modalidades de trabajo en una única plataforma para que las instituciones educativas de la provincia, junto con los actores del sistema educativo y de la administración central, puedan utilizarla. Esta concentración de información permite contar una base de datos unificada, veraz, actualizada en tiempo real y segura, que abarca todas las escuelas de la provincia, así como también a las personas involucradas en el sistema educativo (docentes, no docentes, familia y estudiantes), alcanzando a todos los niveles y modalidades del mismo.

La plataforma GEM registra información administrativa y pedagógica de todas las escuelas de la provincia, incluyendo las de gestión estatal, privada y social. Además, condensa información sobre las y los estudiantes desde que ingresan al nivel inicial hasta que egresan del nivel secundario. También se registran las trayectorias de los niveles o modalidades no obligatorias, como Servicios Educativos de Origen Social (SEOS), Educación Artística o Educación Superior no Universitaria. A esto se suma el registro de todo el personal educativo y de las personas adultas responsables de las y los estudiantes.

En detalle, el registro del GEM está compuesto por 564.664 estudiantes, 3.432 establecimientos, 65.638 personas con cargos docentes y no docentes, y 629.189 madres/padres/tutores, lo que asegura una cobertura provincial del 100%¹.

1 El GEM también funciona como herramienta de administración del personal educativo, registrando toda la información relativa a temas laborales de personal docente y no docente, así como también su recorrido dentro del sistema educativo (condición laboral, licencias, liquidación de sueldos, DD JJ, concursos, bono de puntaje, etc.)

ESQUEMA 2. ALGUNAS DE LAS VARIABLES QUE INTEGRA EL GEM²



Como se mencionó anteriormente, la implementación del GEM dentro del Sistema de Educación de la provincia brinda las herramientas necesarias para el diseño, monitoreo y evaluación de las políticas educativas implementadas en base a los ejes de gestión, así como los insumos para el diseño y funcionamiento del Sistema de Alertas Temprana destinado a abordar la problemática del abandono y la exclusión escolar.

En la siguiente sección, se presentan algunas experiencias de la gestión que contribuyeron a fortalecer el trabajo en red y a incorporar la utilización de los datos en los distintos ámbitos del funcionamiento escolar, esta-

² Listado no exhaustivo.

bleciendo una base sólida para la implementación de otros procesos de innovación en la escuela secundaria, en particular la introducción del SAT.

El uso de los datos en la política educativa. La importancia de focalizar en las trayectorias en riesgo

Promover la inclusión y la equidad educativa es un objetivo primordial que atraviesa todas las acciones de la Dirección General de Escuelas. Está evidenciado que para alcanzar mayor equidad educativa, es imprescindible promover políticas que pongan en el centro del sistema educativo a las y los más vulnerables (Veleda, Rivas y Mezzadra, 2011).

A nivel internacional, el Objetivo de Desarrollo Sustentable (ODS) número 4 establece la necesidad de “garantizar una educación inclusiva y equitativa de calidad y promover oportunidades de aprendizaje permanente para todos, de aquí a 2030” (UNESCO, 2016). En el marco de este ODS, lograr una mayor equidad implica asegurar que nadie quede rezagada/o o excluida/o del sistema educativo, especialmente las poblaciones más vulnerabilizadas. Además, la gestión escolar en la provincia se encuentra en línea con las disposiciones establecidas en la Ley Educativa Nacional (2006), que promueven la prioridad a los sectores más desfavorecidos de la sociedad a través de políticas, estrategias pedagógicas y la asignación de recursos.

La evaluación para la mejora continua constituye un eje fundamental en la política educativa de la provincia, que reconoce la importancia de conocer en detalle y de manera interrelacionada las características de las trayectorias que se encuentran o podrían encontrarse en riesgo escolar, para poder gestionar dando respuestas focalizadas y orientadas de acuerdo cada caso particular.

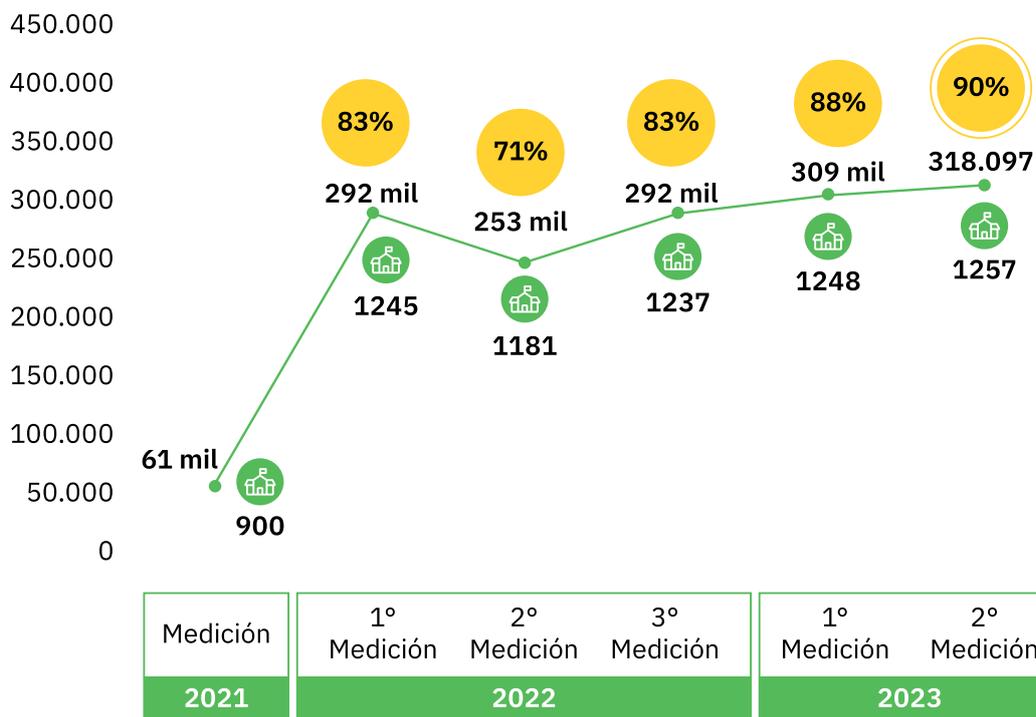
El periodo de pandemia significó un gran desafío para la disminución de la brecha socio-educativa, agravándose

aún más. En este contexto, la Dirección General de Escuelas impulsó una red de apoyo - GEM RED - que se consolidó en base a los datos que el GEM relevó sobre distintas características de las trayectorias de las y los estudiantes. Estas características incluían el nivel de ausentismo de las y los estudiantes, la necesidad de recursos tecnológicos para continuar sus trayectorias y la presencia de dificultades pedagógicas para resolver las actividades escolares.

Esta situación exigió un trabajo mancomunado que involucró el compromiso de las instituciones educativas, equipos directivos y supervisores, equipos docentes y de orientación, entre otros actores. Tanto municipios como institutos de formación docente y universidades, así como las distintas empresas y actores del sector privado, implementaron sus recursos de forma orientada y focalizada de acuerdo a los datos de las trayectorias que arrojaba el GEM. A partir de la disponibilidad y el uso de los datos se logró revincular el 67% de los estudiantes identificados con ausentismo crítico.

En este contexto, entre los años 2021 y 2022, la Dirección General de Escuelas diseñó e implementó una evaluación censal con el propósito de conocer el desempeño de las y los estudiantes en fluidez lectora. Con base en los resultados de esta evaluación, se desarrollaron estrategias de apoyo para aquellos estudiantes que mostraron un rendimiento crítico. Esta evaluación censal de fluidez lectora se enmarca en un camino iniciado en 2017, con un enfoque centrado en la alfabetización como prioridad de la política educativa.

ESQUEMA 3. EVALUACIÓN CENSAL DE FLUIDEZ LECTORA: COBERTURA DE ESTUDIANTES Y ESCUELAS



Fuente: DGE

A partir de los resultados obtenidos en diversas etapas de este monitoreo, en la provincia de Mendoza se consolidó un Plan de Acción Provincial que incluye diversas intervenciones, como talleres de alfabetización y fluidez lectora, la extensión de la jornada escolar y el apoyo pedagógico para estudiantes en situación crítica, entre otras medidas. Además, se estableció un flujo de información y comunicación con las escuelas, proporcionando a las supervisiones y direcciones escolares informes con datos individualizados sobre el rendimiento de cada estudiante.

En 2022, se promulgó la Ley Provincial de Alfabetización (Ley 1.438), sentando las bases para la implementación del Plan Estratégico de Alfabetización (PEAM). Este plan se basa en el uso de evidencia para desarrollar estrategias y programas de alfabetización, así como para realizar un seguimiento individualizado del desempeño de cada estudiante en fluidez lectora. Esto permite moni-

torear los avances y retrocesos en diferentes momentos del año académico.

Además de las políticas centradas en la alfabetización, se ha dado un énfasis significativo a la asistencia escolar, un elemento fundamental en la gestión educativa. Se han implementado diversas acciones destinadas a la reintegración de estudiantes, priorizando aquellas/os con niveles más altos de ausentismo por escuela. A partir de una investigación sobre el ausentismo escolar y sus principales causas que realizó la Dirección de Evaluación de la Calidad Educativa (DECE) de la Provincia de Mendoza junto al investigador Alejandro Ganimian de la Universidad de Nueva York (Ganimian, 2022), se comenzaron a desarrollar acciones para prevenir el ausentismo, teniendo en cuenta distintas dimensiones.

A su vez, en pos de disminuir las brechas socioeducativas, las escuelas con mayores niveles de vulnerabilidad son foco de las políticas de mejora integral de las instituciones. Para esto, desde la DECE, se desarrolló un Índice de Vulnerabilidad Escolar, una medida que sintetiza las condiciones y características de las instituciones educativas de los niveles de escolaridad obligatorios de la provincia, asignando a cada unidad educativa un valor que varía entre 0% (nula vulnerabilidad) y 100% (máxima vulnerabilidad).

Trabajar y gestionar basándose en la evidencia es una modalidad de trabajo que, como cualquier otra, implica un proceso y un período de incorporación. Cuando se considera como una estrategia de gestión institucional, es factible identificar momentos o etapas que se encuentran en continua relación y retroalimentación.

1. El acceso a sistemas de información educativa

Además de la información brindada por GEM, las instituciones cuentan con diversas fuentes de información institucional tales como GEM+³, el Relevamiento Anual⁴ y los operativos de evaluación provinciales, nacionales e internacionales. Así como también, informes periódicos, sistematizados y elaborados desde la DGE, atendiendo las necesidades y particularidades de cada uno de los niveles y modalidades, con datos de relevancia institucional, a saber:

- Informes nominales de Fluidez Lectora.
- Informes con resultados de los operativos de evaluación Nacional Aprender.
- Informes de asistencia de estudiantes.
- Informes de evaluación de programas que responden a ejes prioritarios de gestión.

2. Valoración crítica de las evidencias

Una vez obtenida la información, se debe realizar una valoración crítica de la evidencia, que posibilite transformar percepciones en realidades objetivas, es decir, realizar un diagnóstico complejo de la realidad, identificar fortalezas, debilidades y hacer foco sobre aspectos necesarios y plausibles de modificación. También, este paso permite visibilizar las trayectorias vulnerables que requieren un acompañamiento más cercano, con mayor intensidad.

3. Diseño e implementación de acciones para la mejora

Luego del análisis de las evidencias, es posible diseñar distintas estrategias para fortalecer las trayectorias más debilitadas y establecer líneas de acción orientadas a la

3 GEM+ es una plataforma colaborativa de evaluación de aprendizajes, que permite evaluaciones formativas, así como también, evaluaciones estandarizadas (Relevamiento Provincial de Aprendizajes).

4 El Relevamiento Anual es un operativo educativo llevado a cabo anualmente por la Dirección de Evaluación e Información Educativa del Ministerio de Educación de la Nación. El mismo registra las principales variables del sistema educativo de manera censal.

mejora a partir del diagnóstico realizado, de acuerdo a las distintas herramientas disponibles desde la institución o desde el gobierno escolar, así como desde instituciones locales que articulen con la escuela.

4. Evaluación de las estrategias implementadas

Para conocer en qué medida fue acertado el camino elegido y la efectividad de las herramientas puestas en acción, se propone un análisis desde dos dimensiones. Por un lado, indicadores de proceso que den cuenta del recorrido transitado, para poder tomar decisiones en función del mismo y redireccionar las prácticas, en el caso de ser necesario. Por otro, indicadores de resultados que permitan dimensionar el alcance de los objetivos planteados.

Estas son tan solo algunas de las experiencias y formas de accionar que demuestran el valor del trabajo en red y la importancia de planificar y diseñar acciones detrás de un mismo objetivo que parta de un diagnóstico compartido y claro.

Con todo esto, hay una idea que es importante destacar: contar con datos e información solo es útil si realmente sirven para orientar acciones prácticas y concretas; y tan importante como las mediciones y la información, son las políticas y herramientas que se ponen en juego para dar respuesta a las dificultades identificadas. Y aquí entra en juego otro aspecto clave: el trabajo en equipo con supervisores, equipos directivos y docentes. Los datos permiten identificar la situación de cada trayectoria educativa, pero son los equipos de cada institución quienes conocen de cerca las causas (múltiples y complejas) que están detrás de cada trayectoria en riesgo, y cuál es el mejor modo de acompañar a cada una.

La problemática del abandono y la exclusión escolar

El abandono escolar tiene un impacto significativo en el desarrollo personal, social y económico de las y los jóve-

nes, agravando las brechas de desigualdad en la sociedad. Por tanto, es imprescindible implementar políticas educativas que acompañen las trayectorias de las y los estudiantes y garanticen un abordaje efectivo para quienes se encuentran en situaciones críticas que podrían llevar a la interrupción de sus estudios. Para lograrlo, resulta fundamental contar con información actualizada y de calidad que permita identificar los grupos más propensos a discontinuar sus trayectorias escolares, de modo que puedan diseñarse estrategias de intervención adecuadas.

Actualmente, la educación secundaria representa uno de los mayores desafíos para la política educativa y demanda atención prioritaria. A pesar de haber avanzado sustantivamente en el nivel de acceso y cobertura, aún queda pendiente garantizar el egreso efectivo de todas/os las/os estudiantes, con los aprendizajes necesarios para construir su proyecto de vida. En efecto, el nivel secundario es un punto crítico en el que se observa un aumento tanto en el abandono escolar como en los factores problemáticos relacionados con este fenómeno.

En Argentina, existen varias fuentes de datos que permiten evidenciar y comprender la magnitud de la deserción o exclusión escolar. Algunos de los indicadores que nos ayudan a dimensionar la problemática incluyen la tasa de abandono interanual y de promoción efectiva, el desgranamiento escolar y la tasa promedio de egreso del nivel secundario. Estos insumos, disponibles tanto a partir del Relevamiento Anual y otras bases de la Secretaría de Evaluación e Información Educativa (SEIE) del Ministerio de Educación de la Nación como del Sistema de Gestión Educativa de la provincia de Mendoza, nos permiten analizar las características de la problemática en cada año escolar de cada nivel educativo, así como las variaciones que se producen año tras año.

Al analizar la tasa de abandono interanual en la educación secundaria a nivel nacional, la cual representa el porcentaje de estudiantes matriculados en un año que

no se vuelve a matricular al año lectivo siguiente, se observa una tendencia a la disminución desde el año en que se registró el porcentaje más alto, en el periodo 2016-2017 (10,5%). Para el periodo 2020-2021, se logró reducir significativamente la tasa, alcanzando un 4,5%. Sin embargo, es importante destacar que los datos más recientes muestran un repunte en la tasa de abandono que venía en descenso. Entre los años 2021 y 2022, la tasa de abandono se situó en un 7,6%.

Este mismo fenómeno se evidencia en la provincia de Mendoza. La tasa más alta de los últimos 10 años se registró en el periodo 2014-2015, y desde entonces comenzó un periodo de reducción hasta alcanzar el nivel más bajo en el 2020-2021 (4,7%). Sin embargo, como ocurre a nivel nacional, luego del periodo de pandemia por Covid-19, la tasa de abandono interanual aumentó significativamente llegando al 9,1% (Relevamiento Anual, RedFIE-DIE). Tanto a nivel nacional como provincial, el incremento de la tasa se explica principalmente por un aumento del abandono en 3º, 4º y 5º año entre 2021 y 2022 (entre 5 y 7 puntos porcentuales).

En cuanto a la tasa de promoción efectiva⁵, que indica la cantidad de estudiantes que se matriculan nuevamente en el año inmediatamente superior tras haber promovido oportunamente dicho año, se ha experimentado un aumento significativo en el 2020, alcanzando el 94% a nivel nacional (un incremento de 14 puntos porcentuales con respecto a años anteriores). Sin embargo, en el 2021, luego del contexto de pandemia, la tasa de promoción efectiva cayó 7 puntos porcentuales, llegando al 86%.

En el caso específico de Mendoza, se observa una tendencia similar. La tasa de promoción efectiva tuvo un incremento gradual hasta alcanzar su punto más alto en el 2020 con el 95%, pero sufrió una fuerte caída en el 2021,

⁵ Altos niveles de promoción efectiva señalan bajos niveles de repitencia, reinscripción y abandono (Sistema Nacional de Indicadores Educativos; Ministerio de Educación, Ciencia y Tecnología)

bajando al 85,5% (Relevamiento Anual, RedFIE-DIE). Así como ocurre con la tasa de abandono interanual, los últimos tres años del secundario son los cursos en los cuales se registra una mayor caída de esta tasa.

Respecto a la tasa promedio de egreso del nivel secundario, según el informe de Kit et. al., (2022), de la cantidad de estudiantes que ingresó al primer grado del nivel primario en el 2009, sólo el 53% llegó en tiempo teórico⁶ al último año del nivel secundario. En el caso de Mendoza, ese porcentaje fue del 54%.

Los datos del Sistema de Gestión Educativa de Mendoza (GEM) nos permiten actualizar estos últimos porcentajes. Al analizar el egreso efectivo del nivel secundario a partir de la última cohorte de estudiantes que finalizó la educación secundaria, se observa que de la cantidad de estudiantes que se matricularon en primer año del secundario en el 2018, el 67% llegó a 5° año en el 2022 y, de ese porcentaje, el 55% terminó la trayectoria teórica en tiempo y forma (es decir, el 37% de quienes ingresaron en el 2018).

Al considerar los factores que influyen en la problemática de la interrupción de trayectorias y, por ende, en las causas que llevan a que un/a estudiante esté en riesgo de abandonar y finalmente quede excluido/a del sistema educativo, se encuentran una variedad de factores. Sea una causa única o una multiplicidad de eventos los que debiliten la trayectoria de las y los estudiantes, lo cierto es que antes de llegar al punto de abandonar la escuela, existe un recorrido anterior en el que aún se puede actuar y brindar apoyo para prevenir la exclusión.

En la investigación realizada por la Dirección de Evaluación de la Calidad Educativa - DGE- en el marco de la Investigación Educativa Federal 2022, se buscó contribuir a la comprensión del fenómeno de la deserción y

⁶ 17 años en 5° año del nivel secundario, en el 2016.

exclusión escolar, poniendo el foco en la identificación de factores asociados a la interrupción de las trayectorias escolares. Para lograrlo, se realizaron análisis estadísticos y de regresión con el objetivo de identificar los elementos clave que influyen en esta problemática. Los resultados del estudio, basados en el comportamiento diferencial de variables entre grupos de estudiantes que continúan sus trayectorias y grupos que las interrumpen, evidenciaron hallazgos significativos.

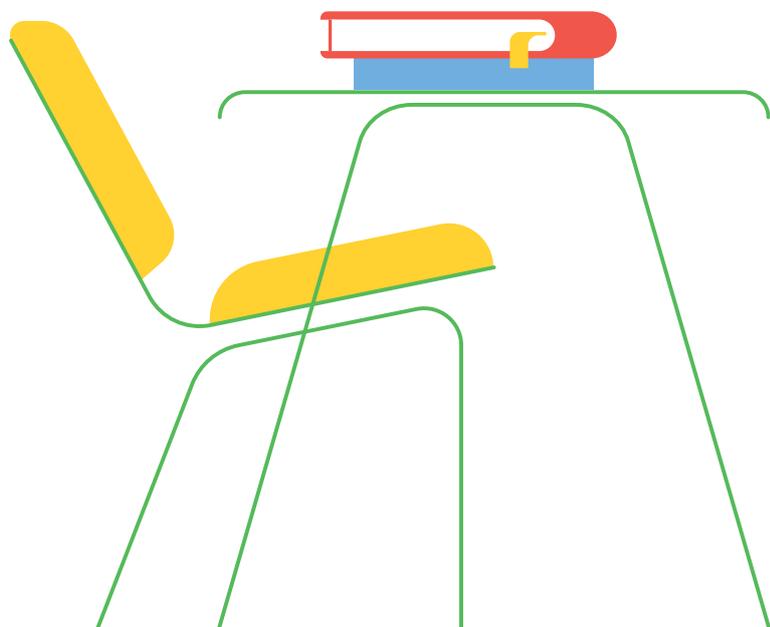
En particular, se encontró que el alto nivel de inasistencias, la sobreedad, el bajo desempeño en materias clave como Lengua y Matemáticas, así como también el desempeño crítico en fluidez lectora y el bajo clima educativo del hogar, actúan -con mayor o menor influencia- como factores asociados al fenómeno de la interrupción de trayectorias escolares.

Aquellos resultados destacan la necesidad de adoptar un enfoque amplio e integral para abordar este fenómeno complejo. En este sentido, es crucial considerar las diversas variables sociales, académicas, económicas y culturales que, de manera interrelacionada, pueden afectar al derecho de la educación.

Es por esto que la disponibilidad de datos confiables y actualizados es fundamental para una gestión educativa efectiva. La utilización de información nominalizada permite comprender patrones de abandono y exclusión escolar, identificar riesgos tempranamente y adaptar políticas e intervenciones a las necesidades específicas de las y los estudiantes. La implementación de un sistema de alerta temprana basado en inteligencia artificial, impulsado por la potencialidad del sistema GEM, representa una herramienta poderosa para prevenir la exclusión escolar y fortalecer las trayectorias más vulnerables.

En el siguiente capítulo, se presentará en detalle el Sistema de Alerta Temprana (SAT) diseñado para prevenir la interrupción de trayectorias escolares. Se explicará

en qué consiste el SAT y cómo se construyó el modelo, aprovechando la información del sistema de gestión nominal de nuestra provincia, GEM. La integración de tecnología y datos precisos permitirá una anticipación más efectiva y una atención personalizada para cada estudiante, garantizando así su educación en tiempo y forma. Descubriremos cómo esta iniciativa se convierte en un paso crucial hacia una educación más inclusiva y exitosa para todas/os.



¿Qué es un SAT y cuál es su utilidad en la prevención del abandono?

¿Qué es un SAT?

Un SAT es una herramienta de apoyo a la política educativa que permite anticipar oportunamente la ocurrencia de un evento no deseado, como la interrupción de una trayectoria educativa. Esta identificación temprana de estudiantes “en riesgo” se nutre de sistemas de información y permite poner en marcha intervenciones específicas para prevenir la ocurrencia del evento (por ejemplo, el abandono escolar).

El SAT proporciona una medida del riesgo de abandono escolar, pero no una certeza absoluta de que el evento efectivamente ocurrirá. Como tal, la medida de riesgo puede asimilarse a una medida de (falta de) solidez de la trayectoria educativa. Cuanto mayor la medida de riesgo, más frágil resulta esa trayectoria y mayores chances existen de que se concrete el abandono en ausencia de intervenciones específicas para atenderlo.

¿Cómo funciona y para qué sirve?

El SAT toma un conjunto de datos vinculados a los factores individuales y contextuales que funcionan como mediadores de la trayectoria escolar y, a partir del análisis y la aplicación de métodos de predicción, estima una medida resumen de “riesgo” para cada estudiante. De esta forma, el SAT genera una lista ordenada de estudiantes desde los de mayor a los de menor riesgo. Esto permite organizar y priorizar intervenciones y contribuye a focalizar los esfuerzos, lo cual es crucial en contextos de recursos limitados y múltiples prioridades.

Una de las principales ventajas de los SAT es justamente su capacidad de resumir los numerosos factores que inciden en el abandono escolar en una sola medida de “riesgo”, lo que facilita la priorización de la atención a los y las estudiantes que más lo necesitan. Así, resultan una herramienta que complementa el conocimiento experto que, en general, ya tienen los y las docentes y el personal escolar, pero que resulta difícil de resumir y, por tanto, complejo de utilizar para establecer prioridades.

Si bien los SAT usualmente emiten alertas a nivel individual (lo que permite priorizar intervenciones desde la escuela) también facilitan la agregación y ordenamiento de información a nivel de grados, escuelas o regiones, lo que puede ser utilizado a nivel de los tomadores de decisión de política educativa para orientar los recursos destinados a políticas y programas preventivos. De esta forma, los SAT permiten escalar de forma rápida y consistente la información que pueden tener docentes y equipos escolares sobre la magnitud del riesgo de abandono escolar en sus propios entornos, y generar insumos para la toma de decisiones a nivel de los ministerios de educación o autoridades regionales correspondientes.

Los SAT son generalmente utilizados para anticipar la interrupción de trayectorias educativas; no obstante, la medida de riesgo podría referirse a otro tipo de evento no deseado. Por ejemplo, en el caso de estudiantes de educación primaria, la medida de “riesgo” más relevante puede no ser el abandono escolar, sino más bien el bajo desempeño académico o la repetición del grado. En definitiva, es posible organizar el SAT de acuerdo con el evento cuya mitigación se desee priorizar.

¿Qué factores considerar en un SAT?

Como ya fue mencionado anteriormente, el abandono escolar es un fenómeno de carácter complejo, que generalmente no es resultado de una única causa sino de un proceso mediado por múltiples factores. Dichos factores

incluyen características individuales (por ejemplo, edad, desempeño académico, si trabaja o no, si tiene menores a cargo); características familiares y del hogar (educación de los padres, ingresos del hogar); y características de la escuela y del entorno. Esos múltiples factores intervienen a lo largo de la vida de un estudiante y durante toda su trayectoria educativa, y pueden constituir señales o síntomas del riesgo de abandonar los estudios (véase Recuadro 1). Los SAT se basan en la presencia de esos factores específicos, individuales y contextuales, que anticipan la interrupción de la trayectoria escolar.

Varios estudios que han explorado la identificación de estudiantes en riesgo de abandono han empleado una combinación de factores para anticipar dicho evento. Entre los más utilizados se encuentran aquellos vinculados a la trayectoria estudiantil, tales como inasistencias, rendimiento académico y evaluaciones de conducta (Adelman et al., 2018; Bowers et al., 2012). Factores asociados al nivel socioeconómico del hogar o características sociodemográficas de la familia como el nivel educativo de los padres están presentes también en algunos de los sistemas de alerta en países de la región. Otros factores que se incluyen en algunos casos son el embarazo adolescente o dedicación a tareas de cuidado (Uruguay, Chile, Colombia, El Salvador), condiciones contextuales en la escuela como clima escolar (Colombia, El Salvador, Costa Rica), o condiciones que indiquen manifestaciones de violencia en el entorno como registro de abusos en el hogar, violencia de género o consumo problemático de sustancias psicoactivas (Belcher & Hatley, 1994; Bowers et al., 2012; Perusia y Cardini, 2021).

Si bien la probabilidad de interrupción de la trayectoria escolar puede asociarse con características individuales, escolares, familiares y comunitarias de los y las estudiantes en riesgo, establecer vínculos causales solo como resultado de un análisis predictivo basado en información estadística puede llevar a conclusiones erróneas. Vale recordar que los SAT proporcionan una

medida de riesgo, pero no informan sobre las causas específicas. Haciendo una analogía, es como si los SAT fueran un termómetro, que toma la temperatura e indica si la persona tiene o no fiebre (y qué tan alta es), pero no indica cuál es la enfermedad específica que está ocasionando la fiebre. Por ello, resulta fundamental complementar esta información con el conocimiento y la capacidad de docentes y personal de las escuelas para indagar y establecer las causas específicas para los estudiantes identificados, que es lo que determinará el tipo de intervención a aplicar.

En definitiva, el papel central de la predicción de probabilidad en un SAT es el de focalizar individuos o grupos con alta probabilidad de abandono, mientras que la identificación de factores causantes de ese potencial abandono debe construirse con información más amplia, que idealmente incorpore el conocimiento que tienen el equipo directivo y el cuerpo docente y no docente de cada escuela.

Recuadro 1. ¿Podemos establecer las causas de abandono desde un sistema predictivo de abandono escolar?

Muchas veces se plantea la posibilidad de ir un poco más allá en el análisis de datos y utilizarlos para determinar los factores que se asocian a una alta probabilidad de abandono. Por ejemplo, puede hacerse un análisis correlacional que indique si determinados atributos de estudiantes/familias/escuelas/comunidad se mueven en la misma (o contraria) dirección a la probabilidad de abandono. Sin embargo, este tipo de análisis puede resultar peligroso por dos

razones: 1) puede confundir las verdaderas causas detrás del abandono escolar; y 2) puede contribuir a estigmatizar a ciertos grupos de población para quienes el riesgo de desvinculación es generalmente mayor.

Esto sucede debido a que la construcción de una medida de riesgo implica una tarea de predicción que básicamente busca en los datos ciertos patrones que se repiten para las personas que suelen abandonar la escuela

y que no están tan presentes en las personas que no abandonan. Estos patrones pueden reflejar las verdaderas causas (sobre las que sería necesario intervenir para poder prevenir), pero también pueden ser solo factores que anteceden en el tiempo al abandono o que meramente reflejan una asociación con las verdaderas causas no observables desde los datos disponibles para el SAT. Por ejemplo, un factor asociado al abandono puede ser la mayor frecuencia de inasistencias. Sin embargo, esto no significa que esas inasistencias sean las causantes de la decisión de desvinculación de la escuela, sino que solo preceden en el tiempo al momento de no asistir más a clases. De manera similar, un factor que se asocia a la probabilidad de abandonar puede ser la zona de residencia, por ejemplo, la residencia en un barrio informal. Esto no significa que residir

en este tipo de barrios sea de por sí una causa de desvinculación, sino que es probable que los residentes en barrios informales cuenten con menores recursos económicos o tengan problemas para movilizarse dentro de las ciudades y sean esas las razones causantes de fondo.

Para poder identificar las variables que causan el abandono, la información estadística necesaria requiere ciertas condiciones (en términos técnicos, exogeneidad de las variables de interés) que solo se dan en contextos muy específicos (por ejemplo, en condiciones experimentales o cuasi-experimentales) que no son las que se tienen cuando se realiza un análisis predictivo como el que surge de un SAT.

Fuente: elaboración propia.

¿Qué se precisa para generar un SAT?

Los SAT precisan, en primer lugar, nutrirse de **datos periódicos y consistentes** a nivel de estudiante, para todos los estudiantes del sistema educativo. Idealmente, para garantizar la sostenibilidad a largo plazo de estos sistemas, se deberían utilizar principalmente registros administrativos y datos nominalizados como fuentes de información. Para ello, los sistemas de información para la gestión educativa (SIGED), organizados para recoger

información nominalizada de las trayectorias estudiantiles, constituyen una herramienta apropiada. Otras fuentes de datos potencialmente útiles provienen de registros administrativos de otras instituciones (salud, desarrollo social) que pueden aportar información relevante sobre el estudiante y/o su contexto. En algunas oportunidades, también puede ser relevante el uso de información que se recolecta de manera específica y complementa a la ya existente.

Dado que generalmente los SIGED están todavía en etapa de consolidación en la mayoría de los países de la región, y también en la mayoría de las provincias argentinas, es importante establecer algunos criterios mínimos de cobertura y calidad sobre los datos, ya que esos atributos afectan decisivamente la calidad de la predicción del riesgo. Los datos que típicamente incorpora un SAT deben ser relevantes para el problema bajo estudio y capturar aspectos individuales, familiares, escolares y comunitarios. Algunos de estos datos variarán con cierta frecuencia (por ejemplo, inasistencias y notas varían con mayor frecuencia que los demás factores), y esta variación da la posibilidad de brindar alertas con frecuencia inferior a la anual. Esto resulta muy importante para reorientar esfuerzos dentro de un año escolar.

Los SAT deben establecer un **método para la estimación del riesgo** de interrupción de la trayectoria. Pueden destacarse tres enfoques: consulta a actores clave, estimaciones basadas en métodos estadísticos tradicionales y predicciones basadas en métodos de aprendizaje automatizado o inteligencia artificial (IA). La consulta a actores clave es el único enfoque posible en ausencia de registros nominalizados que contengan las variables relevantes. Cuando estos registros sí están disponibles, es preferible la aplicación de técnicas estadístico-computacionales, ya sea tradicionales o basadas en IA, dado el mayor grado de acierto en la detección de riesgos. La evidencia disponible que compara estos enfoques indica además que los basados en IA y que incluyen una gran

variedad de variables en la predicción tienen un mejor desempeño predictivo que modelos estadísticos más parsimoniosos (Sansone, 2019).

La elección del método de predicción requiere tener en cuenta varios aspectos. Un primer aspecto importante es la disponibilidad de información. Los modelos basados en conocimiento experto utilizan unos pocos indicadores, a diferencia de los métodos de aprendizaje automático que necesitan un conjunto más extenso de datos, incluyendo información histórica para el entrenamiento del modelo. Un segundo elemento a considerar es la interpretabilidad del modelo, es decir cuán difícil es para los y las docentes y personal de la escuela comprender cómo el modelo toma decisiones y qué variables o características influyen en esas decisiones. Un modelo interpretable aumenta la confianza en el sistema, ya que los usuarios pueden entender el razonamiento detrás de las alertas generadas por el modelo. Más allá de la complejidad del modelo, es fundamental implementar capacitaciones sobre cómo funciona el sistema, qué resultados arroja, qué interpretación corresponde dar a esos resultados, y cómo complementarlos con la información que tiene o puede recoger el docente o la escuela para la toma de decisiones. También debe considerarse la capacidad técnica disponible para operar y actualizar el modelo, lo que generalmente resulta un desafío mayor cuanto más complejo es el método utilizado.

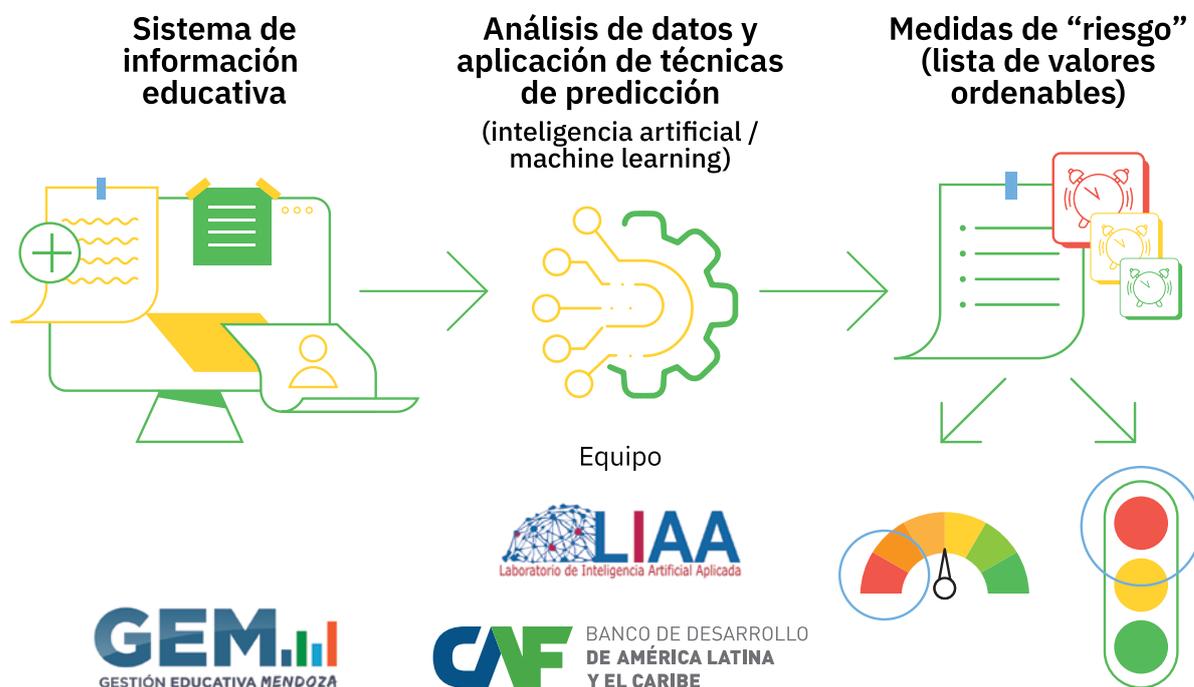
A pesar de estos desafíos, un SAT basado en métodos de aprendizaje automático sobre un buen sistema de información ofrece múltiples posibilidades para informar mejor la asignación de recursos necesaria para atender la desvinculación escolar. Estos métodos pueden tomar grandes volúmenes de información y en general son más precisos, es decir, identifican correctamente a una mayor proporción de estudiantes en riesgo de abandonar los estudios. Asimismo, constituyen modelos flexibles para incorporar o variar los factores considerados,

siendo adaptables a las necesidades de priorización y focalización de los distintos niveles educativos.

El grado de precisión que se puede lograr según el método es un factor relevante a considerar, ya que modelos predictivos imprecisos pueden tener consecuencias no deseadas, tanto para el estudiante como para la escuela o el sistema educativo en su conjunto. Por un lado, la identificación de estudiantes en situación de riesgo que en realidad continuarían su trayectoria sin necesidad de una intervención puede llevar a una asignación ineficiente de los limitados recursos generalmente disponibles para abordar este tipo de intervenciones (Bowers et al., 2012; Gleason & Dynarski, 2002). Por otro lado, el hecho de clasificar a un estudiante como “en riesgo”, especialmente si no lo está, puede ser contraproducente y generar estigmatización, además de impactar negativamente sobre la confianza del estudiante. Sin embargo, el caso inverso resultaría también negativo, ya que si no se detectan casos de riesgo no se tomarán las medidas necesarias de intervención para estudiantes que realmente lo necesitan.

Además de la mayor precisión que caracteriza a los métodos de aprendizaje automático basados en buenas bases de datos, este tipo de sistemas tienen la capacidad de resumir mucha información compleja y entregar fácilmente información utilizable a distinto nivel de toma de decisiones (docente, escuela, ministerio de educación). Finalmente, este tipo de sistemas no necesariamente supone una carga adicional de trabajo a los y las docentes, ya que puede tomar como insumo la información que ellos y otros actores del sistema educativo generan regularmente y que forman parte de los registros disponibles.

FIGURA 1. ¿CÓMO SE GENERA UN SAT?



Los SAT son una de las etapas en el flujo de información para la toma de decisiones. Una vez que son identificados y ordenados los y las estudiantes en riesgo hay que **definir cuáles son las intervenciones** más apropiadas para cada caso. Las intervenciones pueden implementarse en forma personalizada, dirigidas directamente a estudiantes en alto riesgo de abandono por causas específicas (por ejemplo, servicios de orientación pedagógica personalizados, contacto con padres o tutores para profundizar sobre las razones detrás de la alerta y asistirlos en caso que requieran apoyo de la comunidad educativa); en forma grupal focalizada (por ejemplo, clases de apoyo a estudiantes con bajo rendimiento); o en forma universal para toda la escuela (por ejemplo, implementar un protocolo anti bullying o mejorar dotación de ciertos recursos escolares) (Perusia y Cardini, 2021). Las **recomendaciones de acciones preventivas** del abandono escolar pueden basarse a partir de los resultados del análisis de datos; no obstante, de acuerdo

a lo discutido en el Recuadro 1, es preferible basar las recomendaciones de acción en el conocimiento experto de actores clave del sistema educativo, dado su mejor capacidad de detectar las verdaderas causas y de no generar etiquetamientos que atenten contra la autoestima de los y las estudiantes.

También se precisa considerar las **estrategias para conectar a los SAT con la toma de decisiones**. Esta conexión debe contemplar quiénes son los destinatarios de la información de los SAT, que pueden ser distintos miembros de la comunidad educativa, incluyendo a padres, docentes, preceptores, directivos e inspectores/supervisores. Además, esas estrategias deben contemplar el tipo de uso que distintos actores darán a la información provista, por ejemplo, teniendo en cuenta que no todos ellos tienen autonomía para decidir sobre las acciones que se recomienda tomar. En todos los casos, las estrategias deben atender ciertos desafíos que hacen a la credibilidad de la información brindada por el SAT (que depende crucialmente de la calidad de los datos primarios y del desempeño del algoritmo de predicción), a la confianza y nivel de coordinación entre los actores del sistema educativo que deberán implementar las acciones recomendadas, y a la disponibilidad efectiva de los recursos necesarios para esa implementación. Adicionalmente, la conexión de la información de los SAT con la toma de decisiones dependerá de que el esfuerzo y tiempo que implica consultar al SAT no supere los beneficios de hacerlo. Es decir, existe también un componente cognitivo-conductual que considerar a la hora de poner en marcha este sistema de alertas, el cual debe ser transparente y parsimonioso, de manera tal que su consulta no implique una carga adicional a la ya abultada cantidad de información que deben barajar los decisores del sistema educativo.

Finalmente, luego de asignar las intervenciones es importante **monitorear el resultado** de todas las acciones para así retroalimentar el sistema y mejorar su efectividad.

FIGURA 2: EL SAT Y EL FLUJO DE INFORMACIÓN PARA LA TOMA DE DECISIONES



¿Cómo ha avanzado la región en la implementación de SAT?

La experiencia con SATs en la región es reciente y la documentación sobre estas iniciativas está todavía dispersa. Una razón que impide una documentación más sistemática es la diferencia en los enfoques metodológicos para construir las medidas de riesgo, como también en el tipo de datos empleados.

En Guatemala se diseñó e implementó en 2017 la Estrategia Nacional para la Transición Exitosa (ENTRE), enfocada en identificar estudiantes en riesgo en la transición entre la educación primaria y la secundaria. La identificación de dichos estudiantes se realiza mediante un modelo predictivo en base a datos administrativos, y esto es acompañado por un conjunto de intervenciones que incluyen la entrega información sobre cómo prevenir el abandono escolar y recordatorios mensuales mediante mensajes a los directores escolares. La evaluación realizada de esta iniciativa muestra que el SAT acompañado por intervenciones específicas reduce entre 4% y 9% la

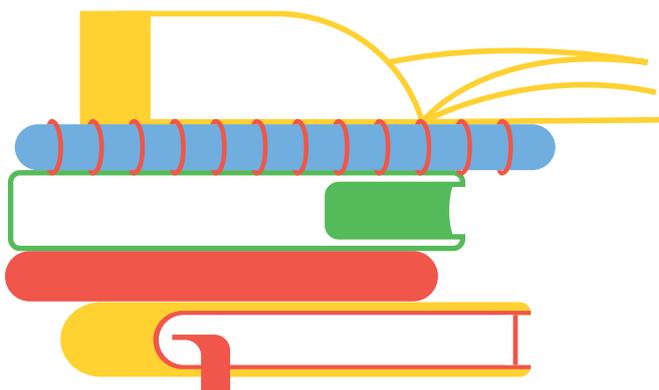
tasa de abandono escolar al finalizar la primaria (Haimovich et al., 2021).

En Chile, el Sistema de Alerta Temprana Escolar se implementó a partir de la pandemia (2020) y fue desarrollado en conjunto con el Ministerio de Desarrollo Social, por lo que incluye factores socioeconómicos además de los propios del sistema de información educativa. El SAT se enfocó de 7mo grado de primaria a 4to año de la secundaria y se implementó inicialmente en un 30% de las escuelas. Tras la identificación de los y las estudiantes, se aplican una serie de estrategias de respuestas individuales para los señalizados en riesgo (visitas, planes curriculares especiales, entrega de material o alimentos) (Juantorena et al., 2022).

En Perú, el sistema de alerta temprana “Alerta Escuela” también comenzó a implementarse en el 2020. El SAT está dirigido principalmente al personal directivo, aunque también tienen distintos niveles de acceso otros actores del sistema educativo. Brinda un sistema de alertas a nivel individual basado en un modelo predictivo que se alimenta de datos de las características del hogar (ingresos, índices de pobreza, planes sociales), del desempeño académico de los y las estudiantes, y del entorno (índices de violencia del barrio).

Más allá de estos ejemplos, en varios países de la región existen distintos niveles de avance en la implementación de SATs. Además de los países ya mencionados, Perusia y Cardini (2021) reportan factores utilizados en sistemas de alerta temprana de Bécice (Early Warning System), Colombia (Sistema de Información para el Monitoreo, la Prevención y el Análisis de la Deserción Escolar), El Salvador (Módulo Complementario del Sistema de Información para la gestión Educativa Salvadoreña), Honduras (Sistema de Alerta y Respuesta Temprana, SART), República Dominicana, Costa Rica (Unidad para la Permanencia, Reincorporación y éxito educativo), Panamá (Red de Prevención y Retención Escolar) y Uruguay (Sistema de

Protección a las Trayectorias Educativas, SPTE). No obstante, se trata en general de experiencias recientes, en proceso de implementación y en general con escasa documentación sobre metodologías utilizadas, niveles de desempeño de los modelos de predicción, o evaluación de resultados e impacto.



Un sistema de alerta temprana basado en el GEM

Como parte de su estrategia de consolidación de la gestión basada en evidencia, la DGE se propuso fortalecer su sistema de alerta temprana de abandono escolar para prevenir la exclusión escolar y fortalecer las trayectorias de los y las estudiantes más vulnerables especialmente del nivel secundario. Gracias a la disponibilidad de datos confiables y actualizados en el sistema de gestión educativa de la provincia (GEM) y la asistencia técnica de CAF-banco de desarrollo de América Latina y el Caribe, fue posible construir un primer modelo predictivo de interrupción de trayectorias escolares basado en IA.

El equipo que desarrolló el prototipo pertenece al Laboratorio de Inteligencia Artificial Aplicada (LIAA)⁷ de la Universidad de Buenos Aires. En enero de 2022, la DGE compartió los datos del GEM con el equipo del LIAA y respondió las consultas que surgieron durante la etapa de preprocesamiento. El equipo del LIAA presentó a la DGE una primera versión del modelo en febrero y una segunda versión en abril. Posteriormente, la DGE compartió datos más actualizados y nuevas fuentes de datos que fueron incorporadas al modelo y dieron lugar a una nueva versión presentada en octubre de 2022. Entre los meses de junio y agosto de ese mismo año, el equipo de LIAA capacitó al personal técnico de la DGE responsable de implementar el modelo desarrollado y brindó apoyo para la estandarización de las consultas a las bases de datos del GEM y del preprocesamiento de los datos. Adicionalmente, capacitaron a funcionarios de la DGE en el desarrollo y la interpretación de modelos de predicción de abandono. Durante todo este proceso, la comunicación entre los equipos de la DGE, LIAA y CAF fue continua.

⁷ El equipo del LIAA firmó un acuerdo de confidencialidad sobre el uso de los datos compartidos. El equipo de CAF no accedió a los datos.

El desarrollo del modelo predictivo fue el primer paso en la consolidación de un sistema de alerta temprana de abandono para la provincia. En las siguientes secciones, se describe el desarrollo del modelo de predicción.

Fuentes de información y variables utilizadas

Los modelos buscan predecir la interrupción de trayectorias escolares (variable de respuesta) a partir de los datos disponibles (variables explicativas). Para medir la interrupción de trayectorias escolares, se utiliza la tasa de abandono interanual⁸. Es decir, para cada estudiante matriculado en un año lectivo el modelo predice si estará matriculado al año lectivo siguiente. Para identificar a los estudiantes matriculados, se utilizan los datos de matriculación nominalizados del GEM para el universo de escuelas de la provincia de Mendoza.

Como se discutió en la sección anterior, los modelos procuran incluir distintas variables explicativas que contemplen factores individuales, familiares, escolares y comunitarios. Gracias a la riqueza de datos del GEM, el modelo desarrollado incluye variables en las cuatro dimensiones (ver Tabla 2). Además de los datos disponibles en el núcleo del GEM, se incorporó al modelo datos complementarios sobre el grado de vinculación de los y las estudiantes con la escuela, la disponibilidad de recursos tecnológicos de los y las estudiantes, el acceso a conectividad y la necesidad de apoyo escolar. La DGE comenzó a recolectar esta información en 2021 a partir de reportes periódicos de los directivos en el marco de GEM-RED.

En cuanto a la frecuencia de actualización de la información en el GEM, la escuela en la que están matriculados los alumnos, sus calificaciones, las inasistencias y la información del GEM-RED se actualizan mensualmente. El

⁸ Se consideró utilizar una variable de respuesta con una granularidad menor definida a partir de las inasistencias, pero se descartó porque es sensible a diferencias en la definición y carga de inasistencias en el GEM (Juantorena et al., 2022).

resto de las variables se actualizan anualmente.

TABLA 1. VARIABLES EXPLICATIVAS INCLUIDAS EN EL MODELO POR TIPO DE FACTOR

Tipo de factor	VARIABLES
Individual	Sexo del alumno Edad Sobriedad Nivel educativo Grado Turno Calificaciones de lengua y matemática: promedio de notas de exámenes intermedios, recuperatorios y trabajos (convertido a terciles dentro de la escuela); cantidad de ausentes en instancias de evaluación; cantidad de evaluaciones compensatorias (recuperatorios) Inasistencias Vinculación con la escuela Requerimientos de apoyo pedagógico Disponibilidad y requerimientos de recursos tecnológicos
Familiar	Si hay datos de los responsables Si tiene tutor Ocupación de la madre y del padre Nivel educativo de la madre y del padre Conectividad en el hogar
Escolar	Tipo de gestión Rural o urbana Modalidad Proporción de varones en la sección
Comunitario	Región, departamento y localidad de la escuela Departamento y localidad de residencia del alumno ⁹ Indicador construido por la DGE que caracteriza la zona de la escuela desde un punto de vista socioeconómico (codificada en deciles) Conectividad en la zona

Fuente: elaboración propia con base en Juantorena et al. (2022).

⁹ Dentro de los datos de los alumnos, se espera poder disponer de una mejor geolocalización del lugar de residencia para poder implementar indicadores de accesibilidad y de caracterización socioeconómica.

Si bien el modelo incorporó las variables señaladas, resulta lo suficientemente flexible como para incorporar nuevas variables a futuro. Por ejemplo, una vez que se tengan suficientes datos históricos de las calificaciones, se podrían incluir en el modelo las notas de los alumnos del año lectivo anterior bajo la hipótesis de que el fenómeno de abandono es el resultado acumulado de varios años y no solo del último año. Adicionalmente, se podría contemplar la inclusión de fuentes de datos externas al GEM que, por distintos motivos, no fue posible integrar en este desarrollo. Por ejemplo, el cuestionario complementario del dispositivo Aprender^{10,11} puede brindar una medida del nivel socioeconómico de los alumnos, y del contexto social de la escuela a partir de indicadores de violencia y convivencia dentro de la escuela. Los datos del Relevamiento Anual de escuelas¹² realizado por el Ministerio de Educación de la Nación permitirían caracterizar a las escuelas en más detalle. Por otro lado, si el censo de fluidez lectora^{13,14} que la provincia realizó en 2021 a alumnos de 4° y 7° grado de la escuela primaria y 1° año de la escuela secundaria se extiende a toda la población escolar, se podrían incorporar al modelo los resultados como marcadores del desempeño escolar individual.

Construcción del modelo de predicción

El modelo construido utiliza herramientas de aprendizaje automático a partir de los datos descriptos del GEM sobre factores individuales, familiares, escolares y comunitarios para predecir durante un ciclo lectivo la falta

10 <https://www.argentina.gob.ar/educacion/evaluacion-informacion-educativa/aprender>

11 El equipo del LIAA accedió a los resultados de las pruebas de lengua y matemática a nivel escuela, pero decidió no incluirlos en la versión final del modelo ya que no hacían un aporte significativo.

12 <https://www.argentina.gob.ar/educacion/evaluacion-e-informacion-educativa/relevamiento-anual-ra>

13 <https://www.mendoza.edu.ar/fluidez-lectora/>

14 El equipo del LIAA accedió a los datos del censo, pero decidió no incluirlos en el presente desarrollo ya que no están definidos para todos los alumnos y solo podrían utilizarse como indicadores a nivel de la escuela.

de matriculación al año siguiente. La carga mensual de los datos de inasistencias, calificaciones y datos del GEM-RED permite una actualización de la predicción mes a mes.

Existen distintos algoritmos de aprendizaje automático que podrían utilizarse en este contexto. Tras un análisis de las características de los algoritmos y de su desempeño en contextos similares, se optó por un algoritmo basado en árboles de decisión llamado *CatBoost* (Prokhorenkova et al., 2018). *CatBoost* tiene la ventaja de que soporta muy bien variables de distinto tipo y datos faltantes, es muy rápido y los valores de los parámetros por defecto proveen un buen punto de partida (Juantorena et al., 2022).

Para que el modelo pueda hacer una predicción sobre datos nuevos, primero hay que entrenarlo utilizando los datos de las variables explicativas y el valor real de la variable de respuesta que se quiere predecir, es decir la tasa de abandono efectiva. Se utilizó la información sobre los alumnos del sistema educativo de Mendoza (variables explicativas) en 2019, 2020 y 2021, y la matriculación (variable de respuesta) en 2020, 2021 y 2022, respectivamente.

Los datos se separaron en cinco subconjuntos preservando las proporciones de estudiantes según departamento, nivel educativo (primario y secundario) y tasa de abandono interanual. Los primeros cuatro subconjuntos de datos se utilizaron para la etapa de entrenamiento. En esta etapa, se probaron distintos modelos con distintas combinaciones de las variables explicativas disponibles y de años para identificar el modelo con el mejor desempeño. Los modelos se entrenaron tomando de a tres subconjuntos de los datos y validando en el subconjunto restante (validación cruzada). El quinto subconjunto de los datos se utilizó para la evaluación final del modelo seleccionado (*held-out*). Es decir, se puso a prueba al modelo con un conjunto de datos que no había sido utilizado para el entrenamiento.

Recuadro 2. Clasificación de estudiantes a partir del modelo, errores y métricas

El modelo predice para cada estudiante la probabilidad¹⁵ de que interrumpa su trayectoria. En la práctica es necesario tomar un umbral de probabilidad a partir del cual se concluye que los estudiantes tienen un alto riesgo de abandonar los estudios. Es decir, estudiantes con una probabilidad menor al umbral son clasificados con bajo riesgo de abandono, y estudiantes con una probabilidad superior al umbral son clasificados con alto riesgo de abandono.

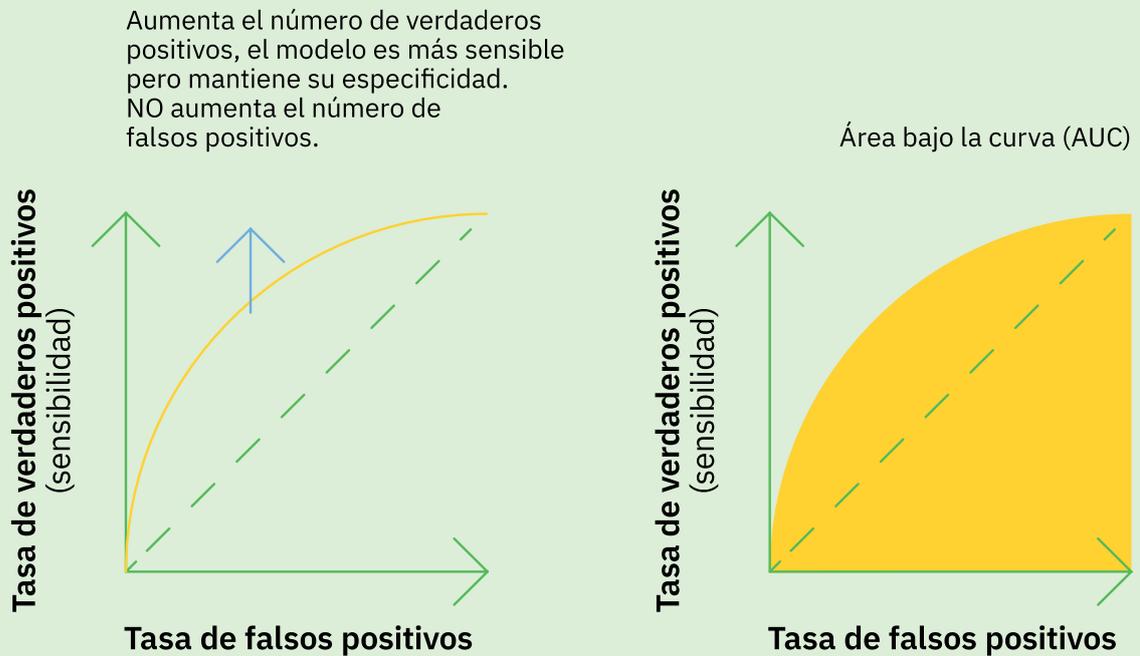
Como se explicó anteriormente, se puede incurrir en errores al clasificar a los y las estudiantes. Uno de estos errores se produce cuando el modelo predice que un estudiante abandonará, pero en realidad no lo hace (error tipo I). El otro error ocurre cuando el modelo predice que un estudiante

no abandonará, pero finalmente sí lo hace (error tipo II). Estos casos reciben el nombre de falsos positivos y falsos negativos, respectivamente.

Para cuantificar estos errores y, por ende, evaluar el desempeño de los modelos de predicción, los estudios utilizan distintas métricas (Bowers et al., 2012). Una de ellas resume el *trade-off* que existe entre minimizar la tasa de falsos positivos y maximizar la sensibilidad o *recall*, definida como la capacidad del modelo de identificar a estudiantes que efectivamente abandonan (a mayor sensibilidad, menor cantidad de falsos negativos). Se grafica la sensibilidad o *recall* versus la tasa de falsos positivos para distintos valores del umbral de probabilidad a partir del cual se considera que un estudiante tiene alto riesgo de abandono, y se calcula el área bajo esa curva (denominada AUC por sus siglas en inglés). Cuanto más cercano a 1 sea el AUC, menor será la tasa de falsos positivos y mayor será la sensibilidad del modelo. Esta métrica permite comparar resultados sin tener que elegir un umbral de probabilidad ya que resume el desempeño del modelo para todos los umbrales.

15 Las probabilidades emitidas por un modelo pueden estar descalibradas, es decir, no corresponder a lo esperado en un caso real. Una probabilidad de 0,5 debería corresponder a que el modelo debería equivocarse la mitad de las veces. Para visualizar si el modelo está calibrado, y por lo tanto poder utilizar sus probabilidades de modo informativo, graficamos la distribución de probabilidad de clasificar a un alumno con alto riesgo de abandono con la proporción de estudiantes que efectivamente abandonan y verificamos que el modelo está razonablemente calibrado.

**FIGURA 3. MÉTRICA DE EVALUACIÓN DE MODELOS:
ÁREA BAJO LA CURVA**



Fuente: elaboración propia con base en Juantorena et al. (2022).

Para comparar el desempeño de los distintos modelos durante la etapa de desarrollo, se utilizan métricas que reflejan la importancia de minimizar los errores de clasificación de los alumnos (ver Recuadro 2). A partir de los valores del AUC¹⁶, se concluyó que el modelo que combina los datos descriptos en la sección anterior provenientes del GEM sobre factores individuales, familiares, escolares y comunitarios, incluyendo datos del GEM-RED, tiene el mejor desempeño. Para decidir qué años utilizar para entrenar el modelo, se comparó el desempeño de un modelo entrenado solo con datos de 2019,

16 El AUC en el conjunto de validación para el modelo entrenado en 2021 con las variables explicativas descritas en la sección anterior alcanza un valor de 0,896.

otro solo con datos de 2021, y un modelo entrenado con ambos años. Si bien el desempeño de los tres modelos resulta similar, el entrenado en 2021 es apenas superior. Además, presenta la ventaja de poder incluir datos del GEM-RED, solo disponible a partir de 2021, que mejora el desempeño del modelo. Los modelos que incluyen los datos de 2020 para el entrenamiento tienen un peor desempeño ya que, dado el contexto de la pandemia, el comportamiento de la tasa de abandono en ese año no es comparable al de otros años. Por lo tanto, se decidió descartar esos datos para la construcción del modelo.

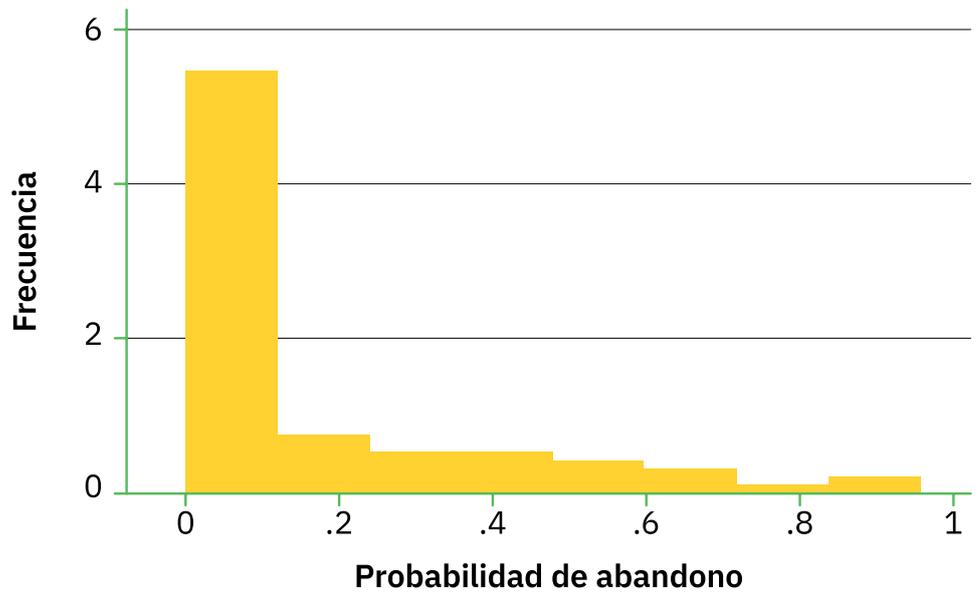
Una vez seleccionado el modelo con el mejor desempeño en la etapa de entrenamiento, evaluamos su desempeño en el subconjunto de datos del *held-out*. En comparación con otras experiencias, el desempeño es muy satisfactorio: el valor del AUC supera el 0,85 y va mejorando mes a mes al tener más información de las calificaciones y las inasistencias¹⁷.

17 El AUC para la predicción de alumnos de nivel secundario es superior a 0,85 para todos los meses, mientras que el AUC para la predicción del nivel primario supera el 0,85 a partir del mes de mayo.

Predicción del riesgo de abandono según características de los alumnos

El modelo genera una medida de riesgo para cada estudiante (ver Recuadro 2). Como se puede observar en la Figura 4, gran parte de los estudiantes tiene una probabilidad de abandono menor a 10%. Sin embargo, como es de esperar, el riesgo de abandono varía según las características de los y las estudiantes.

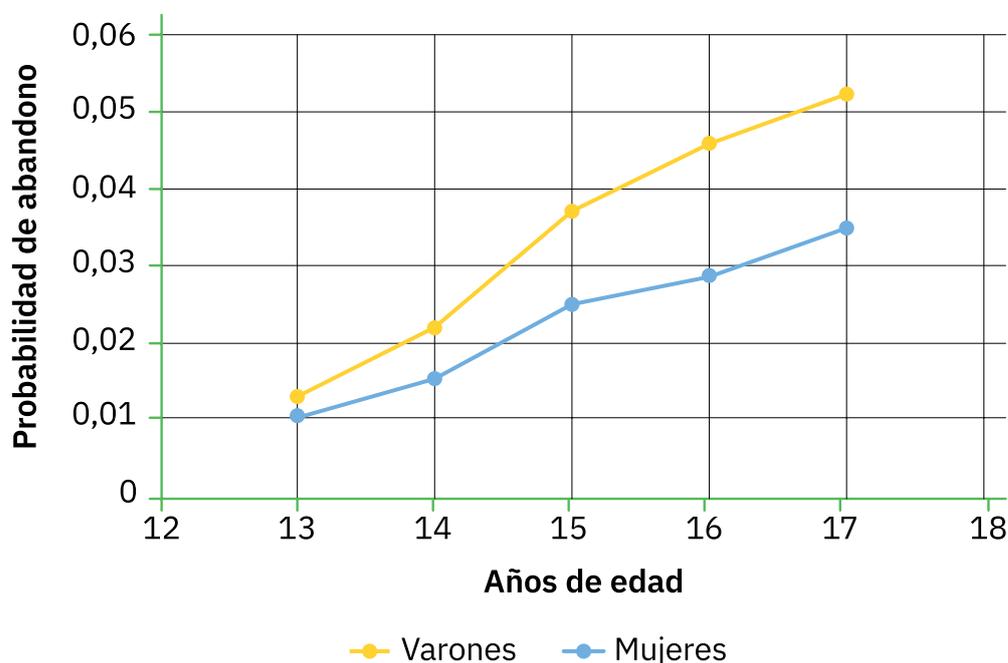
FIGURA 4. DISTRIBUCIÓN DE LA MEDIDA DE RIESGO ESTIMADA POR EL MODELO SELECCIONADO



Fuente: elaboración propia con base en (Juantorena et al., 2022).

Por ejemplo, el riesgo de abandono aumenta con la edad, y lo hace de manera diferente para varones y mujeres (ver Figura 5). A los 13 años, la probabilidad de abandono de los varones es apenas superior a la de las mujeres, pero la brecha se va ampliando entre los 14 y los 17 años.

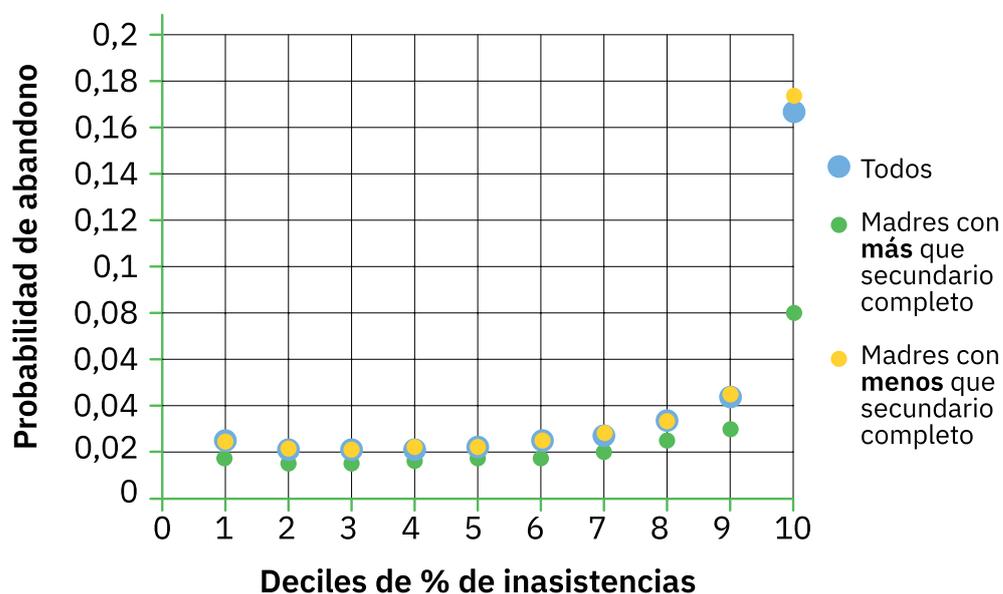
FIGURA 5. PROMEDIO DE LA MEDIDA DE RIESGO POR EDAD Y SEXO



Fuente: elaboración propia con base en (Juantorena et al., 2022).

El riesgo de abandono también aumenta con la cantidad de inasistencias. Adicionalmente, entre aquellos alumnos con mayor cantidad de inasistencias (decil 10), la probabilidad de abandono es mucho mayor entre aquellos con madres con menor nivel educativo (ver Figura 6). También aparecen diferencias notables al comparar a estudiantes de acuerdo con su desempeño académico (notas en matemática/lengua en relación a sus pares), los datos del GEM-RED, la sobreedad, entre otras variables de interés. A su vez, se pueden observar diferencias por departamento, localidad e incluso entre escuelas de una misma zona. Vale aclarar, como se discutió en el Recuadro 1, que esto no implica que estas características sean las causas del abandono escolar, sino que están relacionadas con este fenómeno.

FIGURA 6. PROMEDIO DE LA MEDIDA DE RIESGO POR DECILES DE INASISTENCIAS Y NIVEL EDUCATIVO DE LA MADRE



Fuente: elaboración propia con base en (Juantorena et al., 2022).

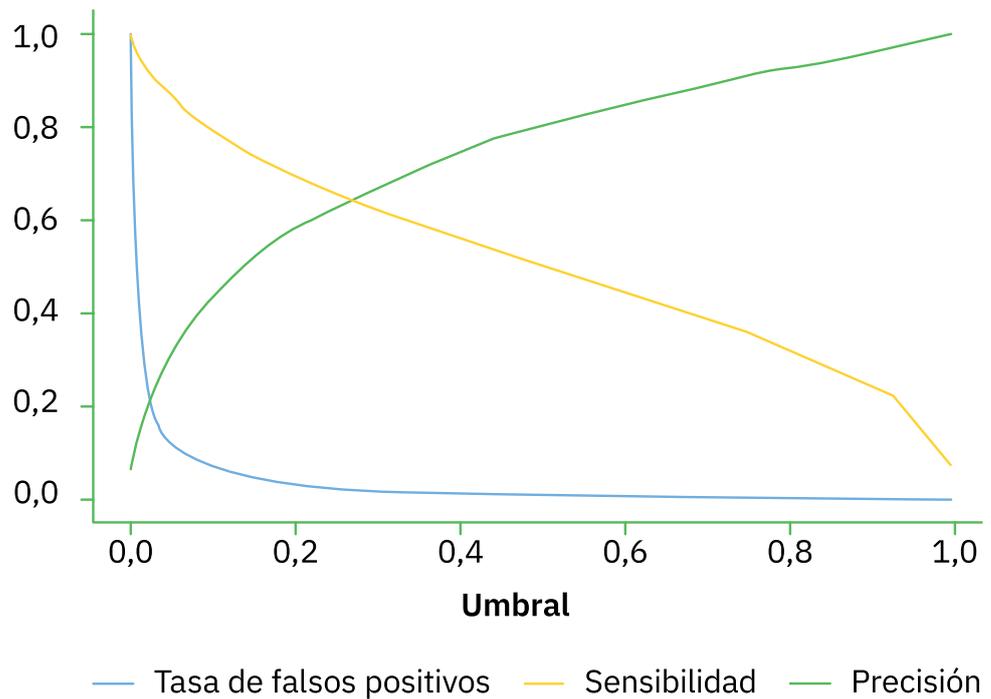
Construcción de un semáforo a partir de la medida de riesgo

Como se menciona en el Recuadro 2, en la práctica es necesario seleccionar un umbral de la medida de riesgo a partir del cual se predice que los y las estudiantes van a interrumpir su trayectoria. Para elegir entre distintos umbrales, se suele comparar la precisión del modelo, que es la tasa de alumnos que abandonan entre los clasificados con alto riesgo de abandono, con la sensibilidad o *recall*.

Cuando se selecciona como umbral un valor pequeño de la medida de riesgo, se clasifican gran parte de estudiantes como potenciales desertores y, por ende, todos los estudiantes que efectivamente abandonan son identificados. Es decir, la sensibilidad del modelo aumenta. Sin embargo, la precisión disminuye porque muchos

estudiantes clasificados como posibles desertores no abandonarán los estudios. Para umbrales más altos aumenta la precisión, pero a costas de no poder detectar estudiantes que abandonan efectivamente. La Figura 7 ilustra esta dinámica y muestra cómo la elección del umbral influye en la sensibilidad y la precisión del modelo

FIGURA 7. MÉTRICAS DE EVALUACIÓN DEL MODELO POR UMBRAL DE LA MEDIDA DE RIESGO



Fuente: Juantorena et al. (2022).

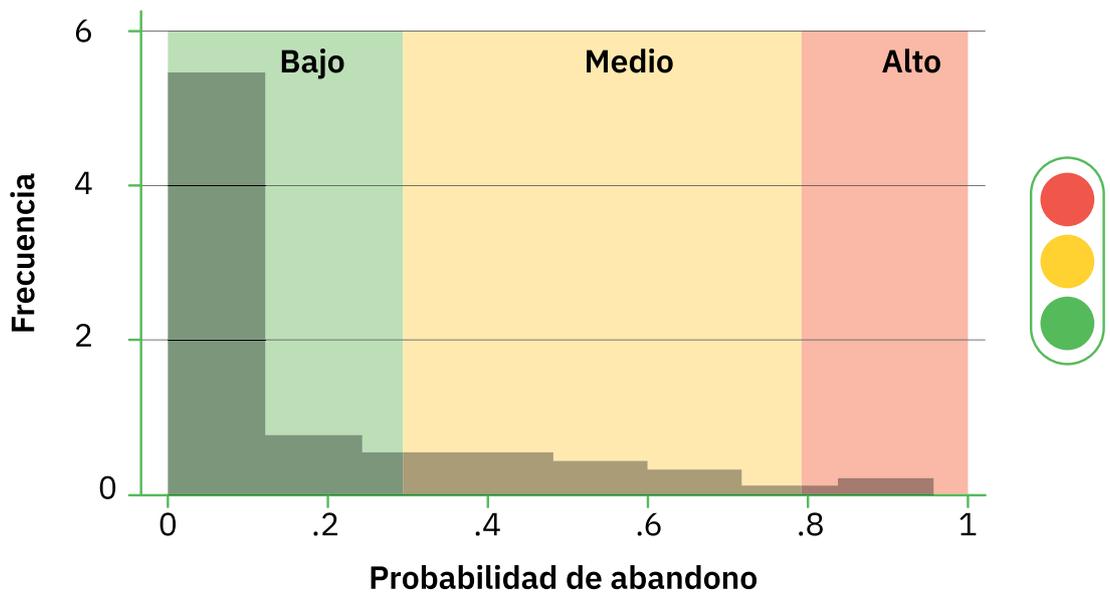
Por lo tanto, para seleccionar el umbral es esencial reflexionar sobre los costos asociados a los errores del modelo, teniendo en cuenta el propósito de las alertas y las intervenciones que se pretenden implementar. Por ejemplo, si las alertas se emplean para asignar recursos, sería deseable que el modelo fuera más sensible para evitar asignar recursos a estudiantes que no corren riesgo de abandonar. Por otro lado, si se considera que las intervenciones dirigidas a estudiantes con alto riesgo de abandono podrían estigmatizarlos, sería importante

mejorar la precisión del modelo. Una opción es asignar costos a las detecciones erróneas: ¿cuál es el costo de no detectar a un estudiante que abandona en comparación con el de tratar a un estudiante que no abandonará como posible desertor? A partir de estos costos, se podría determinar el valor del umbral.

También se debe tener en cuenta cuál es la capacidad para abordar situaciones de riesgo en términos de recursos financieros y humanos. Si se opta por un umbral más bajo para aumentar la precisión del modelo, se debe garantizar que existan los recursos para atender a todos los estudiantes.

En este caso se analizó la posibilidad de utilizar dos umbrales que definen tres regiones o rangos de probabilidad: estudiantes con bajo, medio y alto riesgo deserción (ver Figura 8). Este semáforo permite dosificar los distintos tipos de intervención y mejorar el desempeño del modelo (Juantorena et al., 2022).

FIGURA 8. EJEMPLO DE CONSTRUCCIÓN DE SEMÁFORO A PARTIR DE MEDIDA DE RIESGO



Fuente: elaboración propia con base en (Juantorena et al., 2022).

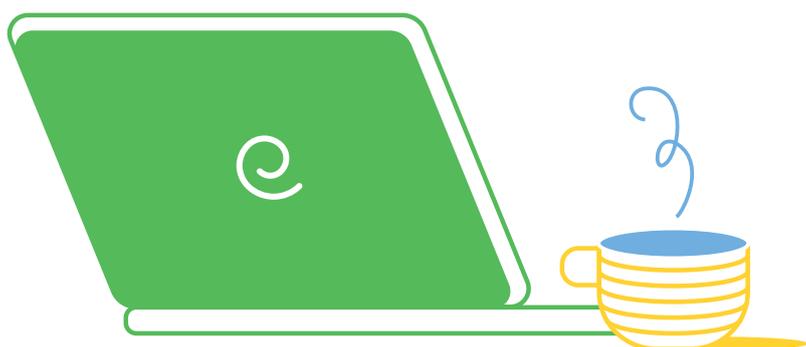
Por ejemplo, para una posible selección de umbrales, el 2% de los alumnos clasificados con bajo riesgo de deserción interrumpen sus trayectorias (falsos negativos), mientras que el 9% de alumnos serían erróneamente categorizados con alto riesgo de deserción (falsos positivos). En este caso, la precisión del modelo es del 91% y la sensibilidad o *recall* es del 24,8%.

TABLA 2. DISTRIBUCIÓN DE LOS ALUMNOS SEGÚN LA PREDICCIÓN DE RIESGO DE ABANDONO PARA EL TOTAL DE ESTUDIANTES

	Bajo	Medio	Alto
Regular	373.552 (98%)	3.207 (50%)	358 (9%)
Deserción	7.979 (2%)	3.133 (50%)	3.665 (91%)

Notas: Se utilizaron umbrales de 0,29 y 0,77 para definir los intervalos, y se estima el semáforo sobre toda la muestra (N = 391.894).

Fuente: Juantorena et al. (2022).



Oportunidades y desafíos para diseñar e implementar SAT

El diseño e implementación de SAT basados en métodos de aprendizaje automático permiten aprovechar la riqueza de datos disponibles en los sistemas de información educativa, como el GEM de la provincia de Mendoza, para abordar la problemática del riesgo de abandono escolar.

En primer lugar, posibilitan una **detección más temprana del riesgo de abandono**, ya que los algoritmos de aprendizaje automático están diseñados para analizar grandes conjuntos de datos de manera eficiente que permiten implementar herramientas al alcance del sistema educativo de manera oportuna. En el caso de Mendoza, la implementación del SAT se llevó a cabo con el propósito de permitir tanto a las escuelas como al gobierno escolar brindar un acompañamiento más intensivo, anticipado y eficiente a las trayectorias reales de las y los estudiantes, mediante acciones destinadas a fortalecerlas y potenciarlas. El modelo emplea datos que se actualizan a lo largo del ciclo lectivo para anticipar posibles interrupciones en las trayectorias educativas durante el próximo año lectivo.

En segundo lugar, estos sistemas permiten construir una medida de riesgo **utilizando toda la información disponible**. Los modelos de aprendizaje automático se “entrenan”, permitiendo que el algoritmo vaya “aprendiendo” a medida que se incorpora nueva información, por lo que constituyen herramientas flexibles para incorporar nuevos datos en el futuro. La incorporación de nueva información permite **mejorar la precisión de la predicción**. En el caso de Mendoza, el modelo hace uso de la información disponible en el GEM y de otros dispositivos que permiten incluir una amplia variedad de factores. El mo-

delo desarrollado es suficientemente flexible para continuar incorporando nueva información que fortalezca aún más la predicción.

En tercer lugar, estos modelos establecen **una única medida resumen** obtenida a partir de toda la información disponible, lo que facilita establecer una priorización de las situaciones a atender y escalar a todos los niveles. En el caso de Mendoza, a partir de las alertas se genera un reporte destinado a equipos de dirección y supervisión, que condensa la información de cada estudiante. La información a la cual se puede acceder a partir de las alertas generadas por SAT incluye la distribución del riesgo de abandono por escuela y la cantidad de estudiantes con riesgo alto y medio por curso y división. Además, permite tanto la identificación por parte de la institución de las causas que están interviniendo en el riesgo de cada trayectoria, como el registro de las intervenciones definidas para acompañarlas y fortalecerlas. Por último, permite evidenciar la cantidad de estudiantes con alertas rojas que fueron foco de intervención, y aquellos que aún no fueron atendidos.

Más allá de sus potencialidades, estos tipos de modelos no están exentos de desafíos. Uno de ellos refiere al propio **manejo del modelo y a su estabilidad** ante los cambios en la generación de información, así como también a **la calidad de los datos** que lo alimentan. Como fue señalado, el sistema se alimenta periódicamente de datos que son preprocesados para entrenar y correr el modelo para generar predicciones regulares. Errores de entrada, duplicación de registros y/o falta de actualización pueden afectar la calidad de los datos y consecuentemente, la capacidad del modelo de identificar correctamente a estudiantes en riesgo. Es necesario que existan capacidades técnicas dentro de los ministerios de educación para realizar un seguimiento y análisis, reportar problemas y generar soluciones para lo cual resulta esencial la capacitación del personal y la implementación de controles de calidad.

En el caso de Mendoza, es relevante señalar que el diseño del modelo fue respaldado por una capacitación ofrecida por el equipo del LIAA al personal técnico de la Dirección de Evaluación de la Calidad Educativa de la DGE, quienes son los encargados de llevar a cabo su implementación. Además, se brindó un apoyo continuo y se resolvieron consultas durante las etapas de carga, procesamiento y análisis de los datos. Asimismo, la formación de funcionarios de la DGE en el desarrollo y la interpretación del modelo resultó ser un elemento fundamental.

Otro aspecto a considerar refiere a la **decisión del valor de los umbrales** para clasificar el nivel de riesgo de interrupción de la trayectoria educativa. Como se expuso anteriormente, la decisión del valor de los umbrales debe considerar los errores de exclusión e inclusión, y las consecuencias no deseadas en la experiencia del estudiante relacionadas por la etiqueta de riesgo. Como se discutió anteriormente, si se toma un umbral más alto la precisión será mayor (menor tasa de falsos positivos), pero también habrá una mayor cantidad de casos de estudiantes no detectados para los cuales no se tomarán acciones. La decisión del umbral también debe considerar la capacidad del sistema educativo de asignar recursos para las y los estudiantes dentro del umbral. Es importante comunicar claramente los umbrales elegidos y las implicaciones asociadas a las partes interesadas. Se puede adoptar un enfoque de umbral dinámico que se adapte a las condiciones cambiantes o a la evolución del modelo con el tiempo.

En el caso de Mendoza, el diseño flexible de los umbrales permitió la calibración del SAT basándose en las experiencias de las escuelas y las alertas generadas. En una primera fase, se implementaron umbrales “piloto” que permitieron una primera identificación de estudiantes con riesgo medio y alto. En una segunda fase, se reentrenó el modelo y se revisaron los umbrales. Esta selección se podrá seguir refinando a medida que se avance en la implementación del sistema.

Un riesgo a considerar para el caso de modelos basados en inteligencia artificial es el de **interpretabilidad**. En general, existe un *trade off* entre la simplicidad y transparencia del método a utilizar y su precisión. Para los equipos docentes, directivos y la comunidad escolar resultará más comprensible un indicador de riesgo de abandono escolar basado en reglas simples y en pocas variables que un método de predicción basado en algoritmos de aprendizaje automático. Para evitar que estos métodos sean percibidos como “cajas negras”, es crucial explicarle a los actores involucrados cómo el modelo ha llegado a sus conclusiones y qué variables o características han influido en la predicción. Un modelo percibido como poco transparente puede impactar negativamente sobre la credibilidad de la información generada y obstaculizar su uso efectivo para una intervención oportuna.

Efectivamente, un riesgo es el bajo **take-up por parte de la comunidad educativa**, que puede percibir esta herramienta como una sobrecarga de trabajo e información. En este sentido, resulta fundamental, por un lado, simplificar y reducir al máximo la carga de información adicional. En este aspecto, el modelo implementado en Mendoza utiliza información que ya está cargada en el GEM por lo que no representa una sobrecarga de trabajo en las escuelas para generar las alertas. Por otro lado, se precisa incorporar a la comunidad escolar en la conexión con la toma de decisiones sobre el paquete de medidas para cuidar las trayectorias de las y los estudiantes. La percepción de que el sistema produce información que permite generar acciones es lo que promoverá su uso efectivo.

Con el propósito de mitigar los efectos de estos riesgos (baja interpretabilidad, bajo take up), una medida fundamental es la capacitación a todas las personas involucradas en el manejo de los datos, para entenderlos y usarlos, así como la comunicación efectiva sobre la interpretación de la medida de riesgo y el involucramiento de la comunidad escolar en la toma de decisiones de intervención.

En la experiencia de implementación y la internacionalización de la herramienta en la gestión escolar diaria en la Provincia de Mendoza, se ha destacado la participación organizada de la comunidad escolar. Esta participación activa involucra a los equipos de supervisión, dirección y docentes. La gestión escolar ha trabajado en la construcción de procedimientos tanto para el registro y uso de la información como para abordar las trayectorias identificadas en riesgo. Puntualmente, las causas de riesgo que identifican los equipos directivos, así como las intervenciones que se planifican, se encuentran sistematizadas, categorizadas y piloteadas.

El proceso de adopción de esta herramienta en Mendoza, iniciado a mediados de 2022, incluyó múltiples fases de sensibilización y preparación con diversos actores internos y externos al sistema educativo. La prueba piloto del SAT implicó capacitaciones a funcionarios, equipos técnicos y equipos escolares. El monitoreo y la sistematización de los resultados del piloto permitieron la elaboración de recomendaciones y el diseño de un plan de comunicación de resultados pertinente y útil para los principales actores involucrados.

Durante el año 2023, se diseñó un plan de acompañamiento territorial a las escuelas por parte de la Dirección General de Escuelas, con el objetivo de registrar información sobre sus experiencias. Se promovieron distintas instancias de revisión y puesta en común de propuestas pedagógicas para abordar las trayectorias en riesgo, identificando los programas, planes y proyectos ejecutados por las escuelas, así como los recursos humanos y financieros dispuestos. Todas estas acciones contribuyeron a la apropiación adecuada y responsable por parte de los tomadores de decisiones a distintos niveles, al involucramiento de los actores de la comunidad educativa y a la concientización sobre el uso de la herramienta.

Finalmente, un desafío importante para la implementación de estos sistemas es su **escalabilidad y sosteni-**

bilidad; es decir, la factibilidad de que el sistema esté disponible en todas las escuelas, se alimente, se use, dispere intervenciones para mitigar el riesgo de abandono escolar y se sostenga en el tiempo. En el caso de la provincia de Mendoza, el sistema se encuentra actualmente operativo en todas las escuelas secundarias logrando una cobertura del total provincial. Para el año 2023, el porcentaje de riesgo de interrupción de trayectorias educativas fue calculado en dos momentos (abril y noviembre) para aproximadamente 150.000 estudiantes de nivel secundario, utilizando los datos obtenidos del GEM y el modelo de predicción.

En paralelo, la provincia implementó, con el apoyo de CI-PPEC, una iniciativa piloto en 17 escuelas secundarias, para la cual se conformó un equipo de acompañamiento compuesto por integrantes de la Dirección de Evaluación de la Calidad Educativa, Dirección de Planificación de la Calidad Educativa, Dirección de Educación Secundaria y Dirección de Educación Técnica y Trabajo. La implementación de esta iniciativa piloto permitirá realizar una evaluación del primer año de implementación e identificar desafíos y eventuales ajustes en el proceso, tales como necesidades adicionales de capacitación y/o acompañamiento por parte de supervisores, ajustes a la plataforma de visualización, revisión de umbrales de riesgo, frecuencia de la emisión de alertas, o abanico de intervenciones a poner a disposición para los centros educativos.

Algunos **desafíos a futuro** para la mejora continua incluyen:

- Asegurar la calidad de los datos y ampliar a nuevas fuentes de información, que permita garantizar mayor precisión y confiabilidad, enriquecer la base de datos y mejorar la capacidad predictiva.
- Incorporar datos longitudinales, con el fin de mejorar la comprensión de las trayectorias y su evolución en el tiempo.
- Consolidar la incorporación de la herramienta en cada una de las instituciones, a través del apoyo ins-

titucional necesario para maximizar su impacto.

- Generar evaluaciones sistemáticas de las acciones e intervenciones implementadas como respuesta a las alertas, para mejorar su capacidad de abordaje de las trayectorias en riesgo.
- Continuar trabajando en la apropiación del modelo y su reentrenamiento anual e incorporar nueva información por parte de los equipos técnicos.

Sin desmedro de los desafíos pendientes, el Sistema de Alertas Tempranas de la provincia de Mendoza está construido sobre bases sólidas: datos de un sistema de información robusto como el GEM y excelentes estándares científicos para el modelado de la predicción. Si bien resulta de una aplicación de inteligencia artificial, ha sido implementado y usado por diferentes actores del sistema educativo. Su uso consistente y responsable puede permitir importantes ganancias de retención de estudiantes en el sistema educativo y contribuir, de ese modo, a mejorar las oportunidades educativas de las y los adolescentes y jóvenes de la provincia.



Referencias

Adelman, M., Haimovich, F., Ham, A., & Vazquez, E. (2018). Predicting school dropout with administrative data: New evidence from Guatemala and Honduras. *Education Economics*, 26(4), 356–372. <https://doi.org/10.1080/09645292.2018.1433127>

Belcher, D. C., & Hatley, R. V. (1994). A Dropout Prediction Model That Highlights Middle Level Variables. *Research in Middle Level Education*, 17(2), 67–78. <https://doi.org/10.1080/10825541.1994.11670032>

Berniell, L., Díaz, B., Estrada, R., Hatrick, A., Llambí, C., Maris, L., & Singer, D. (2021, August 17). Políticas para reducir las brechas educativas en la pospandemia. Caracas: CAF. Retrieved from <https://scioteca.caf.com/handle/123456789/1773>

Berniell, L., de la Mata, D., Bernal, R., Camacho, A., Barrera-Osorio, F., Álvarez, F., ... Vargas, J. (2016). RED 2016. Más habilidades para el trabajo y la vida: los aportes de la familia, la escuela, el entorno y el mundo laboral. Bogotá: CAF. Retrieved from <https://scioteca.caf.com/handle/123456789/936>

Bianchi, B., Pietto, M. L., & Kamienkowski, J. E. (2019). *Estimación de la Interrupción de las Trayectorias Escolares en Escuelas Secundarias Públicas de la Provincia de Buenos Aires*. Programa Manos en la DATA 2018. Educación.

Bowers, A. J., Sprott, R., & Taff, S. A. (2012). Do We Know Who Will Drop Out? A Review of the Predictors of Dropping out of High School: Precision, Sensitivity, and Specificity. *The High School Journal*, 96(2), 77–100.

Ganimian, A. J. (2022) “Informe sobre las Ausencias de los Estudiantes a la Escuela en la Provincia de Mendoza, Argentina. Incidencia, Antecedentes y Posibles Consecuencias.” Nueva York, NY: Universidad de Nueva York.

Gleason, P., & Dynarski, M. (2002). Do We Know Whom to Serve? Issues in Using Risk Factors to Identify Dropouts. *Journal of Education for Students Placed at Risk (JESPAR)*, 7(1), 25–41. https://doi.org/10.1207/S15327671ESPR0701_3

Haimovich, F., Vazquez, E. J., & Adelman, M. (2021). Scalable early warning systems for school dropout prevention: Evidence from a 4.000-school randomized controlled trial. *Documentos de Trabajo Del CEDLAS*, no. 285. <http://sedici.unlp.edu.ar/handle/10915/121906>

Juantorena, G. E., Kamienkowski, J. E., y Riera, P. E. “Reporte técnico final: Predicción de Interrupción de Trayectorias Escolares para un Sistema Alerta Temprana en la Provincia de Mendoza” (Agosto 2022) Manuscrito no publicado.

Kit, I., España, S., Catri, G., Nistal, M., & Volman, V. (2022). Desgranamiento y aprendizajes desiguales: las dos caras de la misma moneda. *Argentinos por la Educación*, 7.

Perusia, J. C., & Cardini, A. (2021). Sistemas de alerta temprana en la educación secundaria. *Prevenir el abandono escolar en la era del COVID*, 19.

Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., & Gulin, A. (2018). CatBoost: Unbiased boosting with categorical features. *Advances in Neural Information Processing Systems*, 31.

Sansone, D. (2019). Beyond early warning indicators: high school dropout and machine learning. *Oxford bulletin of economics and statistics*, 81(2), 456-485.

UNESCO (2016) “Guía Abreviada de Indicadores de Educación para el ODS 4”

UNESCO, CEPAL y UNICEF (2022): La encrucijada de la educación en América Latina y el Caribe: informe regional de monitoreo ODS4-Educación

Veleda, C. Rivas, A. y Mezzadra, F. (2011): La construcción de la justicia educativa. Criterios de redistribución y reconocimiento para la educación argentina, CIPPECUNICEF-Embajada de Finlandia, Buenos Aires.



