



**UNIVERSIDAD
TORCUATO DI TELLA**

MASTER IN MANAGEMENT + ANALYTICS

**Diseño de políticas de apoyo escolar
dirigidas a estudiantes de primaria
con alto riesgo de repetición: Un
enfoque basado en modelos de
aprendizaje automático**

Elianna Testa Smirne

Mayo 2022

Tutor: Gabriel Martos Venturini

Resumen

El sistema educativo uruguayo se ha posicionado en los últimos años entre los peores de la región en lo que refiere a culminación de la educación media y a la repetición en educación primaria. A pesar de que hay bastante evidencia empírica reflejada en una larga discusión en artículos y revistas especializadas de que estos dos fenómenos están relacionados, la mayor parte de las acciones que se toman para prevenir la desvinculación de la educación formal se destinan a estudiantes del nivel de educación media.

Es por este motivo que el presente trabajo tiene como objetivo, a través de modelos estadísticos y de aprendizaje automático, identificar a los estudiantes de educación primaria con mayor riesgo de repetición, para así poder tomar las acciones preventivas más adecuadas a los fines de evitar la deserción escolar e incidir de forma positiva para que los estudiantes en riesgo logren completar con éxito el ciclo de educación primaria y tengan también mayores chances de terminar el ciclo de educación media y superior.

Abstract

In recent years the Uruguayan education system has positioned itself among the worst in the region in terms of completion of secondary education and repetition in primary education. Despite the fact that there is enough empirical evidence reflected in a long discussion in articles and specialized journals that these two phenomena are related, most of the actions taken to prevent disengagement from formal education are aimed at students from the middle education level.

By the use of statistical models and machine learning this work aims to identify who are these primary school students who have a higher repetition risk, in order to be able to take the most appropriate prescriptive actions. These actions would help prevent school dropout and have a positive impact, so that students at risk could successfully complete the primary education cycle and also have greater possibilities of finishing the secondary and higher education cycle.

Índice

1. Introducción	5
1.1. Contexto	5
1.2. Problema	8
1.3. Antecedentes	10
1.4. Objetivo	12
2. Datos	13
2.1. Descripción de datos recolectados	13
2.2. Análisis descriptivo	17
3. Metodología	18
3.1. Ingeniería de atributos	19
3.2. Modelos a explorar	19
3.2.1. <i>Feature selection</i>	19
3.2.2. <i>Random Forest</i>	20
3.2.3. <i>Boosting machines</i>	21
4. Resultados de la modelización y estrategias a aplicar	23
4.1. Resultados de la modelización	23
4.2. Estrategias a aplicar	24
4.2.1 <i>Aprovechamiento de plataformas disponibles</i>	25
4.2.2 <i>Trabajo en conjunto con las comunidades</i>	26
4.2.3 <i>Incorporación del equipo del Sistema de Protección de Trayectorias</i>	26
4.2.4 <i>Evaluaciones y seguimiento</i>	27
4.2.5 <i>Tutorías individuales</i>	27
4.2.6 <i>Participación de la familia</i>	27
4.2.7 <i>Desestructuración del aula y enseñanza a través de lo lúdico</i>	28
4.2.8 <i>Adaptaciones curriculares</i>	28
4.2.9 <i>Educación emocional</i>	28
Conclusiones	30
Bibliografía	32
Apéndice A. Información adicional de variables	34
Jurisdicciones	34
Tipos de centros educativos	34
Apéndice B. Ingeniería de atributos	35
Edad del adulto al momento del nacimiento	35
Faltas	35

Índice de tablas

Tabla 1. Total de variables recolectadas	13
Tabla 2. Matriz de confusión para Random Forests.....	20
Tabla 3. Matriz de confusión para Boosting.....	22
Tabla A 1. Localidades.....	34
Tabla B 1. Faltas justificadas en Y1 según grado	36
Tabla B 2. Faltas justificadas en Y1 según quintil	36
Tabla B 3. Faltas justificadas en Y1 según zona.....	36
Tabla B 4. Faltas justificadas en Y1 según turno.....	37
Tabla B 5. Faltas injustificadas en Y1 según grado	37
Tabla B 6. Faltas injustificadas en Y1 según quintil	38
Tabla B 7. Faltas injustificadas en Y1 según zona.....	38
Tabla B 8. Faltas injustificadas en Y1 según turno.....	38
Tabla B 9. Faltas justificadas en Y2 según grado	39
Tabla B 10. Faltas justificadas en Y2 según quintil	40
Tabla B 11. Faltas justificadas en Y2 según zona.....	40
Tabla B 12. Faltas justificadas en Y2 según turno	40
Tabla B 13. Faltas injustificadas en Y2 según grado	41
Tabla B 14. Faltas injustificadas en Y2 según quintil	41
Tabla B 15. Faltas injustificadas en Y2 según zona.....	41
Tabla B 16. Faltas injustificadas en Y2 según turno.....	42
Tabla B 17. Faltas justificadas en Y3 según grado	43
Tabla B 18. Faltas justificadas en Y3 según quintil	43
Tabla B 19. Faltas justificadas en Y3 según zona.....	43
Tabla B 20. Faltas justificadas en Y3 según turno	43
Tabla B 21. Faltas injustificadas en Y3 según grado	44
Tabla B 22. Faltas injustificadas en Y3 según quintil	45
Tabla B 23. Faltas injustificadas en Y3 según zona.....	45
Tabla B 24. Faltas injustificadas en Y3 según turno.....	45

Índice de figuras

Figura 1. Jóvenes de 20 a 24 años con educación secundaria completa	7
Figura 2. Egreso en educación media superior	8
Figura 3. Asistencia de jóvenes de entre 15 y 17 años.....	8
Figura 4. Estudiantes según zona de la escuela	17
Figura 5. Estudiantes que repitieron según quintil y género	17
Figura 6. Estudiantes con familias beneficiarias de TUS	18
Figura 7. Estudiantes según la situación laboral de la persona a cargo	18
Figura 8. Estudiantes según la edad de la persona a cargo.....	18
Figura 9. Importancia de las variables en Random Rorest	21
Figura 10. Importancia de las variables en Boosting Machines	22
Figura 11. Resultados en Random Forests y Boosting.....	24

1. Introducción

Alcanzar una educación de calidad y generar equidad a través de la misma se ha mantenido como uno de los principales objetivos por parte de todos los gobiernos uruguayos desde el retorno a la democracia. Sin embargo, en el sistema educativo se observan actualmente altos niveles de desvinculación y repetición, tanto en educación primaria como en educación media.

A través de este trabajo se pretende contribuir en tomar acciones sobre los estudiantes que tienen mayor riesgo de fracaso escolar, para atenderlos lo antes posible y generar un impacto dentro de sus trayectorias educativas. Particularmente, se espera que a través de modelos estadísticos y de aprendizaje automático se pueda identificar cuáles son los niños y niñas que tienen más probabilidad de repetir su curso en primaria y así llevar a cabo acciones concretas sobre ellos.

A través del mismo, se podrá observar que con la utilización de distintas variables vinculadas al centro educativo, al contexto familiar y al entorno socioeconómico del estudiante, es posible predecir cuáles son los alumnos que se encuentran en mayor situación de vulnerabilidad, es decir sobre los que hay que tomar acciones por tener un mayor riesgo de repetición.

A continuación, se describen las características del sistema educativo uruguayo y su posicionamiento en relación a otros países de la región, los problemas que se identifican alrededor del fracaso educativo, junto a las tendencias que se observan a nivel internacional sobre el mayor aprovechamiento de los datos y de la tecnología.

En las próximas secciones se presentan los datos con los que se trabajó para el análisis, así como modelos que fueron desarrollados y los resultados a los que se llegó. Finalmente se realizan sugerencias acerca de cómo se puede aprovechar de mejor manera la información generada.

1.1. Contexto

Uruguay es un país de tres millones y medio de habitantes, con una tasa de natalidad que se ha visto disminuida en los últimos años. Tiene una significativa estabilidad democrática, un índice de Gini de 0,38 y el 10,2% de su población se encuentra por debajo de la línea de pobreza, aumentando a 18,5% si se considera solamente a las infancias de entre 6 y 12 años.¹ A su vez, el país enfrenta actualmente grandes desafíos en lo que respecta a su sistema educativo.

La Administración Nacional de Educación Pública (ANEP), es el ente autónomo estatal responsable de la planificación, gestión y administración del sistema educativo público. Se encuentra dividido en cuatro subsistemas: la Dirección General de Educación Inicial y Primaria, la Dirección General de Educación Secundaria, la Dirección General de Educación Técnico Profesional y el Consejo de Formación en Educación.

A su vez, su órgano rector es colegiado, el Consejo Directivo Central (CODICEN), y es designado por el Poder Ejecutivo con venia especial del Senado.

La escolaridad es obligatoria a partir de los 4 años, asistiendo los estudiantes a partir de los seis años a educación primaria, durante seis años. Luego a educación media básica durante tres años

¹Inicio - Instituto Nacional de Estadística. (s/f). Gub.uy. Recuperado el 1 de mayo de 2022, de <https://www.ine.gub.uy/>

y a educación media superior durante otros tres años. Se dice que una trayectoria educativa es exitosa cuando se culmina la educación media superior a los 17 o 18 años.

Existe amplio consenso acerca de la importancia de una educación de calidad para la construcción de sociedades más igualitarias, inclusivas y justas, siendo la educación pública el principal motor para la movilidad social ascendente. José Pedro Varela, el fundador de la escuela laica, gratuita y obligatoria en Uruguay, pregonaba que “para crear el gobierno del pueblo, lo primero es despertar, llamar a la vida activa, al pueblo mismo; para hacer que la opinión pública sea soberana, lo primero es formar la opinión pública; y todas las grandes necesidades de la democracia, todas las exigencias de la república, solo tienen un medio posible de realización: educar.”². Sin embargo, en los últimos años Uruguay ha aumentado su gasto público en educación, mientras que se identifica un importante rezago y abandono, manteniendo bajas tasas de culminación de la educación media y ubicando al país entre los peores del continente.

A su vez, hay evidencia de que las desigualdades observadas al interior del sistema educativo son reproducidas por el mismo sistema. Por ejemplo, la forma en que se asignan los docentes en la ANEP genera una segmentación en la distribución de los mismos, generando que quienes se encuentran más y mejores calificados se ubiquen en centros de contextos socioeconómicos más favorables. A su vez, el mecanismo según el cual se asignan los estudiantes a los centros también se vuelve reproductor de la desigualdad en el sentido de que los alumnos son asignados según la zona en la que viven. Por último, la forma en que se proveen los recursos materiales a los centros es la misma para todo el país, lo que genera que aquellas escuelas que tienen comisiones de padres más fuertes y en mejor situación económica tengan mejor infraestructura, ya que esta es normalmente mejorada por estas comisiones.³

En esta línea, al analizar los indicadores vinculados al sistema educativo uruguayo, tomando valores al 2018, los cuales no están influidos por la pandemia COVID19, se observa que solamente la mitad de los jóvenes de 16 años asisten sin rezago a la educación formal, y si tenemos en cuenta a quienes tienen 19 años, vemos que solo el 36% culminó la educación media superior.⁴

El sistema educativo uruguayo actualmente recibe a 244.303 alumnos en educación primaria (de 1ro a 6to año). El registro histórico de los datos educativos del sistema señala que en promedio el 4,68% de la matrícula anual repite de curso--lo que equivale a aproximadamente 11.500 estudiantes repitentes cada año-. A su vez, en el primer año de escolaridad la incidencia promedio es sensiblemente más alta y repiten en promedio el 11,32% del total de la matrícula. En el total acumulado a lo largo del período escolar (primaria y educación media), aproximadamente el 30% del total del estudiantado del sistema público cursa el mismo año al menos más de una vez. A pesar de que la repetición se ha visto disminuida en los últimos años, se mantiene como la más alta de América Latina⁵

Por otro lado, se destaca que Uruguay se ubica como el país de Latinoamérica con menor tasa de finalización de la educación secundaria.⁶

²Varela, J. P. (1874)

³ Bogliaccini, J. y Rodríguez, F. (2015)

⁴INEEd. (2019)

⁵Aristimuño, A. (2009).

⁶ Aristimuño, A., Giorello, M., y Lasida, J. (2020)

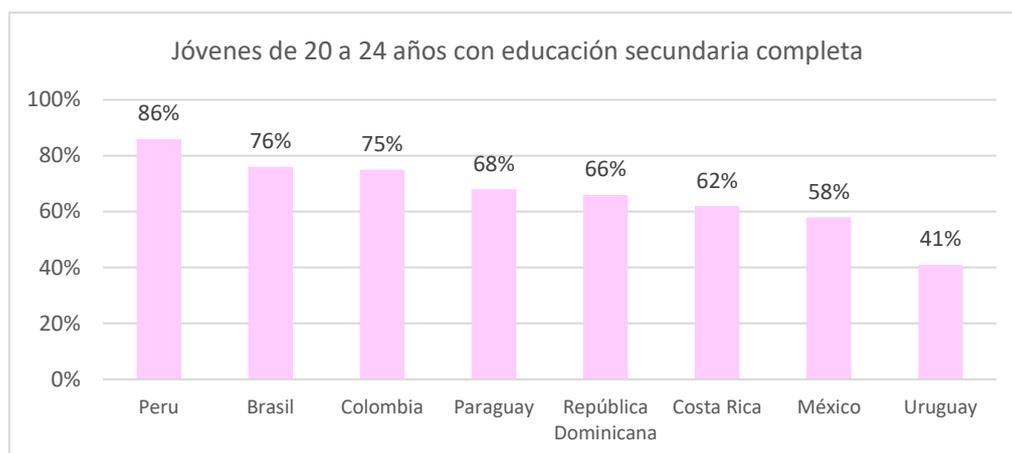


Figura 1. Jóvenes de 20 a 24 años con educación secundaria completa⁷

Estos indicadores dan cuenta de que la juventud uruguaya no logra tener trayectorias estudiantiles continuas y exitosas, siendo a veces estas truncadas o demasiado prolongadas en el tiempo. Estas trayectorias se ven afectadas por la baja asistencia, la repetición y el abandono.

Es necesario destacar que Uruguay ha realizado importantes progresos en lo que se refiere a cobertura educativa, sobre todo en nivel inicial y en primaria. Entre niñas y niños de 6 a 12 años se ha logrado una cobertura prácticamente universal (98% de la población asiste a educación primaria), y también se ha logrado en los últimos años poder ampliar la matrícula en educación media básica. La misma pasó de una tasa neta de cobertura de 67,4% en el año 2000 a 81,1% en el 2019.⁸

Este es un fenómeno que se ha visto durante los últimos años en toda América Latina. “En los años recientes, en paralelo con la democratización del acceso a la educación secundaria, [...] se amplía el refugio para los sectores populares que, históricamente no habían accedido a los estamentos superiores del sistema. Es un proceso de dos vías, que debilita la frontera mientras la establece por primera vez como membrana protectora para quienes logran romper las cadenas intergeneracionales de la exclusión escolar temprana. En buena parte de América Latina, la caída del poder escolar coincide con su redefinición como refugio para los excluidos sociales.”⁹

Sin embargo, a pesar de la ampliación de la cobertura lograda los indicadores vinculados al fracaso escolar no son buenos. Observando las trayectorias educativas vemos que los peores resultados de repetición se dan en educación media, pero numerosos estudios muestran que uno de los factores determinantes para la repetición y luego el abandono en educación media es el haber repetido en educación primaria¹⁰. El trayecto por la educación formal es acumulativo, por eso se pretende poner foco en este estudio en la educación que se brinda entre los 6 y 12 años, para así tomar acciones enfocadas a las infancias y lograr una mejor y más efectiva incidencia sobre las trayectorias educativas de niños, niñas y adolescentes.

A su vez, se observa una importante desigualdad al interior del sistema educativo, ya que los resultados se distribuyen de forma heterogénea según el nivel socioeconómico del

⁷INEEd. (2021)

⁸Inicio – Observatorio de la Educación. (s/f). Edu.uy. Recuperado el 1 de mayo de 2022, de <https://observatorio.anep.edu.uy/>

⁹Rivas, A. (2019)

¹⁰Aristimuño, A. (2009)

estudiantado. Uruguay mantiene la brecha más alta de Latinoamérica en lo que refiere a culminación de la educación secundaria.¹¹

El 18% de jóvenes de quintil 1 que tienen entre 21 y 23 años culminaron en 2018 la educación media superior. Entre quienes pertenecen al quintil 5, la finalizó el 82%.

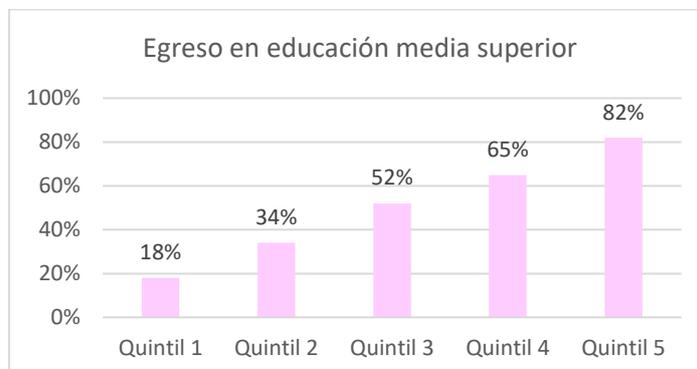


Figura 2. Egreso en educación media superior¹²

En esta misma línea, entre jóvenes de 15 a 17 años, el 74% de quienes tienen ingresos altos asisten al curso que les corresponde según su edad (edad oportuna), mientras que solamente el 27% de quienes tienen ingresos bajos asiste a su curso correspondiente.¹³

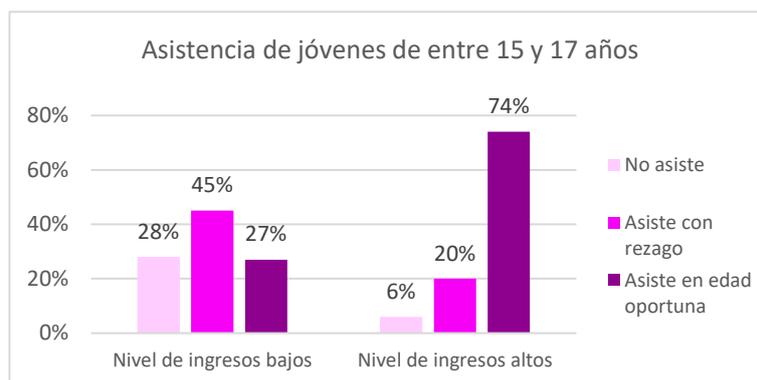


Figura 3. Asistencia de jóvenes de entre 15 y 17 años¹⁴

1.2. Problema

Cuando se habla de fracaso escolar, existe amplio consenso acerca de que es multicausal y que los factores asociados al mismo provienen tanto del sistema educativo como del entorno del estudiante. Es así que en esta sección distinguiremos a los componentes vinculados con la comunidad educativa en la que se enmarca el estudiante de los relacionados con el origen y el contexto en el que este vive y se desarrolla.

En relación al contexto del estudiante, se observa que el fracaso escolar afecta de distinta manera a quienes pertenecen a los sectores socioeconómicos más vulnerables. La evidencia muestra que “mientras más pobres, vulnerables y excluidos son los estudiantes, mayores son sus probabilidades de no aprender lo necesario, de no alcanzar buenos desempeños, de

¹¹Aristimuño, A., Giorello, M., y Lasida, J. (2020)

¹²DIEE-ANEP en base a ECH-INE (2019)

¹³DIEE-ANEP en base a ECH-INE (2019)

¹⁴ DIEE-ANEP en base a ECH-INE (2019)

reprobar grado, de dejar de asistir a clases o finalmente desertar definitivamente del sistema escolar.”¹⁵

A su vez, numerosos estudios vinculados a la Sociología de la Educación dan cuenta de los procesos de selección escolar y de cómo estos no son equitativos, que quienes provienen de entornos socioeconómicos y socioculturales más desfavorecidos “repiten más y tienen peores rendimientos educativos, que tienen menos presencia en la escolarización postobligatoria y que obtienen, finalmente, menos credenciales educativas de las que les corresponderían respecto a su presencia poblacional”¹⁶. Al mismo tiempo, el abandono del sistema educativo genera una mayor tendencia a cometer actos delictivos así como a un consumo abusivo de sustancias o actividades sexuales de riesgo, las cuales representan amenazas a su bienestar futuro.¹⁷

También el nivel educativo de la familia juega un papel importante sobre el éxito de las personas en la educación formal. El capital cultural y simbólico que rodea su ambiente familiar, la manera en que valoran sus madres y padres la educación y cuál es la utilidad que le encuentran a la misma inciden sobre el desarrollo educativo del estudiante. Las expectativas que mantienen sobre sus hijos también juegan un rol. Cuanto más altas son las expectativas que mantienen madres y padres, más alta es la probabilidad de que sus hijos permanezcan en el sistema educativo formal. También en esto influye la trayectoria que la familia y el entorno han tenido en el sistema, si lograron concluir sus estudios o si también desertaron.¹⁸

Cuando hablamos de factores vinculados al sistema educativo, incluimos a las causas intraescolares, que tienen que ver con las situaciones y dinámicas que se observan al interior del sistema, por ejemplo, el traspaso del estudiantado entre la primaria y la educación media, y el hecho de haber repetido uno o más grados. La probabilidad de abandonar el sistema educativo aumenta según más grados haya repetido la persona. Cuanto mayor es la extraedad más se ve incrementado el riesgo de abandono y desvinculación de la escuela y sistema.”¹⁹ Es entonces un fracaso escolar construido desde la escuela.

Por otro lado, han habido otros cambios culturales según los cuales las normas y rituales asociados a la escuela han comenzado a cambiar, vinculadas con cambios en la experiencia escolar y que repercuten en que en ciertos casos los alumnos ya no encuentran buenas razones para seguir estudiando, no solo por el costo de oportunidad de estar fuera del mercado laboral, sino por “la cercanía entre el currículo y las expectativas y posibilidades de los estudiantes o la valoración que los jóvenes hacen de la educación como una herramienta para su desarrollo.”²⁰

A su vez, estos factores se podrían ubicar en dos perspectivas conceptuales “Un primer subgrupo destaca la visión del fracaso escolar como resistencia a los códigos socializadores que entrega la escuela [...]. Así, la escuela en vez de organizarse y ofrecer estrategias adecuadas para que los estudiantes con distinto capital cultural puedan integrarse y asimilar la cultura escolar predominante, reconociendo y validando su propia cultura, los segrega y estigmatiza al asumirlos como sujetos carentes de las capacidades y disposiciones para aprender.”²¹

El segundo subgrupo se vincula principalmente con aspectos que son propios del plantel docente y de su interacción con el estudiantado. En muchos casos se observa que estos tienen

¹⁵Román, M. (2013)

¹⁶Cabrera, L. (2019)

¹⁷Josephson, K., Francis, R., & Jayaram, S. (2018)

¹⁸Román, M. (2013)

¹⁹Román, M. (2013)

²⁰Aristimuño, A. (2009)

²¹Román, M. (2013)

“dificultad para transferir conocimientos y ofrecer a sus estudiantes herramientas y estrategias cognitivas y socio afectivas que les permitan aprender y fortalecer capacidades y habilidades en este proceso”²², otro factor determinante suele recaer en las expectativas que tiene este docente sobre sus estudiantes. Diversos estudios suelen mostrar que una imagen negativa de los docentes acerca de los niños se traduce en un nivel de logros menor y en un peor rendimiento.²³

Dados los resultados observados previamente, se podría decir que el sistema educativo no estaría siendo tan igualador de oportunidades, sino que estaría perpetuando las inequidades que se observan desde el punto de partida. El abandono escolar tiene orígenes claros en la repetición previa, siendo esta la primera de muchas experiencias de fracaso escolar que vivirá el estudiante a lo largo de su vida. Es por esto que se vuelve indispensable atender a estas primeras experiencias de repetición.

Sumado a la discusión, es necesario destacar que más allá de las altas tasas de repitencia que presenta Uruguay actualmente, este mecanismo ha formado históricamente parte del sistema educativo, al igual que lo que sucede en el resto de América Latina. Sin embargo, en países de América del Norte o de Europa Occidental, el instrumento de la repetición se ha eliminado o es insignificante. A pesar de que los países sin reprobación son los que obtienen mejores resultados en las pruebas PISA o IEA,²⁴ no necesariamente la falta de utilización del instrumento de la repetición corresponde con que los alumnos tengan una mejor performance, sino que hay evidencia que da cuenta de los efectos negativos que tiene la repetición sobre las infancias y adolescencias.

Los argumentos a favor de la repetición refieren a una necesidad de alcanzar un mínimo umbral de conocimientos o la falta de maduración de estos estudiantes, así como que puede ser una medida para promover el esfuerzo, pero la evidencia que va en contra de esta medida afirma que cuando se realiza a edades muy tempranas hay una incerteza en la mejora de estos rendimientos, mientras que genera desmotivación entre estudiantes y un desenganche escolar, afectando al logro educativo posterior.²⁵

Además, hay evidencia acerca de cómo se podría sustituir la repetición por otros instrumentos, brindando diferentes tipos de apoyos a estudiantes que no logren los umbrales de competencias correspondientes.²⁶

1.3. Antecedentes

Actualmente la ANEP posee una gran cantidad de información sobre quienes asisten al sistema educativo en su totalidad, la cual se nutre de los cuatro subsistemas entre los que se encuentra dividido. Cada uno de estos gestiona su información a través de sus respectivas direcciones de tecnología y de sus direcciones de estadística y la ponen a disposición de la población en general

²²Román, M. (2013)

²³Román, M. (2013)

²⁴UNICEF. (2012)

²⁵Cabrera, L. (2019)

²⁶UNICEF. (2012)

y del personal de los centros educativos a través de los Monitores Educativos²⁷ de cada uno de estos subsistemas.

A su vez, el CODICEN de la ANEP creó la División de Investigación, Evaluación y Estadística (DIEE), a cargo de la gestión del Observatorio de la Educación²⁸, que publica una evolución de los principales indicadores relativos a estudiantes y centros educativos.

Estos monitores y el observatorio son utilizados en algunos casos por las inspecciones o direcciones de los centros, pero hay una sensación general de que esta información se usa poco para la toma de decisiones.²⁹ Se evidencia que hay una necesidad de mejora en unificar esta información y en el hacérsela llegar a los centros educativos de forma amigable y comprensible, pero también, y principalmente, a los referentes y docentes que están en contacto con las niñas, niños y adolescentes del sistema.

Por otro lado, a nivel internacional se evidencia la posibilidad de una nueva revolución educativa, en el sentido de las transformaciones que se están dando alrededor de las tecnologías de la información y las posibilidades de generación de conocimiento e innovaciones.³⁰ En línea con esto, se ha visto en los últimos años un aumento del uso de machine learning para el área de la educación, donde las aplicaciones van desde la predicción de la performance de los alumnos hasta el uso de técnicas de aprendizaje automático para evaluar y calificar a los estudiantes de manera más rápida y justa.

Además, en lo que se refiere a la aplicación de las tecnologías de la información en la educación, se observa el auge de sistemas de Gestión de Aprendizaje (LMS), que generan un atractivo para los estudiantes y ayuda a generar nuevas experiencias de aprendizaje³¹. También las nuevas tecnologías pueden ayudar a crear tareas personalizadas para cada estudiante, al mismo tiempo que mantienen un mayor control sobre el trabajo que estos van realizando, ayudando a los docentes a tener un mayor y mejor monitoreo de sus estudiantes.

A su vez, otras experiencias ponen a los docentes como objetivo, como por ejemplo las flipped classroom, o las plataformas que los apoyan para la calificación de estudiantes, permitiéndoles así un mejor aprovechamiento de su tiempo. Algunas de estas tecnologías han sido foco de crítica en el sentido que un aumento en la corrección mediante máquinas podría afectar los lazos de poder que los docentes mantienen con sus estudiantes, en relación a las relaciones de admiración o autoridad. Pero como positivo, se valora la retroalimentación y la personalización que se vería sustancialmente mejorada.³²

Se observa entonces el surgimiento de un nuevo mercado tecnoeducativo, que cada vez avanza más y denota como “el nuevo orden mundial de la educación está atravesado por la promesa de la tecnología.”³³

²⁷ *Monitores educativos*. (s/f). Edu.uy. <https://www.anep.edu.uy/estadisticas-evaluaciones/monitores-educativos>

²⁸ *Inicio*. (s/f). Edu.uy. <https://observatorio.anep.edu.uy/>

²⁹ Aristimuño, A., Giorello, M., y Lasida, J. (2020)

³⁰ Brunner, J. J. (2000)

³¹ Rivas, A. (2019)

³² Rivas, A. (2019)

³³ Rivas, A. (2019)

1.4. Objetivo

Más allá de las consecuencias positivas que podría llegar a tener el eliminar la repetición como instrumento dentro del sistema educativo, en el mediano plazo Uruguay no parecería estar abriendo la discusión hacia la supresión de esta herramienta. Es por esto que asumiendo que la repetición es un fenómeno que se va a continuar dando en la educación formal en los próximos años este trabajo pretende, tomando en cuenta factores vinculados al entorno del estudiante, así como algunos relacionados con el propio sistema educativo, contribuir a la disminución de la repetición en educación primaria y así ser una herramienta más en la protección de las trayectorias educativas de los estudiantes.

La razón por la que esta primera aproximación al problema pretende concentrarse sobre la población de estudiantes de primaria es porque no hay dudas de que la deserción se asocia fuertemente con historias de repetición reiterada. Se ha observado en Uruguay que “el 88% de los desertores habría repetido antes y el 70% había repetido más de una vez. Para un alumno que repitió en primaria, la tasa de riesgo de desertar es 4,3 veces mayor que el que no repitió en dicho tramo, con todas las otras condiciones iguales.”³⁴

Es así que tomando acciones sobre los potenciales repitentes de estos grados se podrá impactar de forma más temprana en la mejora en las trayectorias educativas de todos los estudiantes uruguayos. Además, de esta manera no solo prevenimos futuras instancias de repetición, sino también generamos una mejor dinámica en la eficiencia administrativa de los recursos del organismo en los próximos años.

Es sabido que la investigación educativa se enfoca en entender las múltiples causas asociadas a la deserción estudiantil y particularmente a la repetición en educación primaria, y no es el propósito de este trabajo intentar sustituir estas investigaciones, sino el poder realizar un mejor aprovechamiento de los datos disponibles y las tecnologías para contribuir a la toma de decisiones informadas por parte de los centros educativos y del sistema educativo en su conjunto.

El objetivo de este trabajo es entonces modelizar y estimar mediante modelos estadísticos y de aprendizaje automático la probabilidad de que estudiantes de educación primaria (1ro a 6to grado) repitan su próximo curso lectivo, con el objetivo de identificar de forma temprana a los estudiantes más vulnerables. Con los outputs de estos modelos se podrán tomar acciones específicas para cada estudiante con altas probabilidades estimadas de repetir de curso, de manera de minimizar las posibilidades de que esto suceda, al mismo tiempo que se liberarían recursos del presupuesto educativo que podrían ser destinados a otros programas de la ANEP que tiendan a mejorar la calidad y el alcance del sistema de educación público.

El trabajo de esta tesis servirá como prueba piloto para que en el futuro mediato se extrapolen los métodos y acciones que de ella surjan al estudio y prevención de la repitencia en los siguientes niveles de educación del sistema público.

³⁴Román C., M. (2019)

2. Datos.

2.1. Descripción de datos recolectados

Mediante un pedido de utilización de datos al CODICEN de la ANEP se accedió a la información del total de la matrícula de educación primaria durante los años 2021, 2020 y 2019.

Educación primaria utiliza para el registro de todos los asistentes a sus centros educativos el sistema GURI, a partir del cual se puede obtener información acerca de la trayectoria del estudiante en educación primaria pública. Cabe destacar que en casos de educación privada, esa información no está integrada.

A su vez, se utilizaron datos del Ministerio de Desarrollo Social (MIDES), que se recolectan a través del Sistema de Información Integrada del Área Social (SIAS). Este sistema está conformado por la mayoría de las instituciones públicas, y sus cometidos principales están relacionados con “la conformación de un servicio de información social que integre datos y registros administrativos de los organismos responsables de las distintas políticas sociales en Uruguay (Salud y Alimentación, Educación, Soluciones habitacionales, Seguridad y Protección Social, Trabajo y Empleo, Discapacidad y Atención integral).”³⁵

Es así que además de datos históricos acerca del paso del estudiante por el sistema educativo, se pudo incorporar información en relación a la situación laboral de sus padres, de protección social y de soluciones habitacionales.

Se presenta entonces a continuación el listado de variables para el año 2021, que son las mismas que se recabaron para los años 2020 y 2019. A continuación se brinda una descripción breve de cada una de las mismas y luego se detalla si son numéricas (num) o categóricas (cat).

Tabla 1. Total de variables recolectadas

	Variable – Año 2021 ³⁶	Descripción breve	Tipo
Datos del grupo	grado	Grado que cursará cada estudiante en el 2021.	cat
	comedor	Indica 1 si asiste al comedor ubicado en el centro educativo.	cat
	turno	Turno en el que asiste el estudiante al centro.	cat
Condición del estudiante	genero	Género del estudiante.	cat
	extraedad	Indica 1 si tiene extraedad. Esto significa que está por encima de la edad teórica que debería tener al momento de cursar el grado en que lo está haciendo. (Esta variable se calculó a partir de la fecha de nacimiento, tomando en cuenta la fecha límite hasta la cual un estudiante puede ingresar a determinado grado).	cat
Datos del centro	tipoCentro	Tipo de centro al que asiste. En el apéndice se detallan los tipos de centros existentes al año 2021.	cat
	departamento	Departamento donde está ubicado el centro educativo al que asiste.	cat

³⁵ ¿Qué es el SIAS? (s. f.). Ministerio de Desarrollo Social. <https://www.gub.uy/ministerio-desarrollo-social/etiqueta/otros/es-sias>

³⁶ El total de las variables se trataron como integer en el momento de generar los modelos en R

	jurisdiccion	En Uruguay hay 19 departamentos, pero a nivel de inspecciones de educación primaria, Canelones y Montevideo están subdivididos según la ubicación geográfica de los centros, por lo que el país se divide en 24 jurisdicciones. Se detallan también en el apéndice las jurisdicciones.	cat
	localidad	La localidad es una unidad geográfica menor a la jurisdicción, simula algo cercano a los barrios donde se ubican los centros educativos.	cat
	zona	Indica 2 si el centro se ubica en una zona urbana y 3 si se encuentra en una zona rural.	cat
	quintil	Quintil socioeconómico de la escuela. Esta definición la realiza la División de Estadística de CODICEN, en relación a características de las familias que asisten a los centros educativos, así como según la zona donde están ubicados. Se detalla también en el apéndice cómo se calculan estos niveles socioeconómicos.	num
Historial del estudiante	faltas_injustificadas	Es la cantidad de faltas injustificadas que tuvo el estudiante en el año anterior al identificado. Para los datos de 2021 es el total de faltas injustificadas que tuvo durante el 2020.	num
	faltas_justificadas	Total de faltas justificadas que tuvo durante el año escolar anterior.	num
	faltas_totales	Suma total de faltas que tuvo a lo largo del año anterior.	num
	repitio	Indica si el estudiante repitió en el año anterior al analizado. En caso de 2021, indica si repitió en 2020.	cat
	calificacion_conducta	Indica la calificación que tuvo en conducta en el año anterior al analizado. Para el 2021 correspondería la información de 2020, pero por motivo de la pandemia se decidió durante ese año que los alumnos no tengan calificación, solo se indique si debían repetir o no. Es por esto que el dato de calificación conducta está disponible para 2019 y 2018.	num
	calificacion_rendimiento	Indica la calificación en rendimiento. En relación al año escolar 2020, surge la misma limitación que en la calificación de conducta.	num
Protección social	inauClub	Indica si este estudiante luego de la jornada escolar, asiste a una propuesta de club de niños del INAU. ³⁷	cat

³⁷“La misión de INAU es promover, proteger y/o restituir los Derechos de niños, niñas y adolescentes, y garantizar el ejercicio efectivo de su ciudadanía a través de distintas instancias de participación.

Los Clubes de Niños son centros socioeducativos de atención diaria, pensados para complementar la acción de la familia y de la escuela en la educación, socialización, desarrollo, crianza y mejora de la calidad de vida de los niños y niñas. Son también propuestas de ampliación del tiempo pedagógico: incorporan nuevos contenidos educativos y amplían el tiempo de cuidado. La atención se brinda contra horario escolar (turno matutino y/o vespertino), en locales adaptados a las necesidades y características de la franja etaria. Tienen además presencia en todo el país.” Clubes de Niños. (2017, noviembre 22). Gub.uy. <https://www.inau.gub.uy/infancia/clubes-de-ninos>

	inau_acogimientoFamiliar	Muestra si tiene registro en algún programa de la modalidad "acogimiento familiar" ³⁸ de INAU.	cat
	inau_programaCalle	Identifica si tiene registro durante el año en el sistema de respuesta de tiempo parcial para niños en situación de calle de INAU. ³⁹	cat
	mides_refugios	Indica si está asistiendo a centros del MIDES en los que se brinda alimento, alojamiento y abordaje interdisciplinario para la adquisición de hábitos de la vida cotidiana, cuidados de salud y recuperación de redes sociales. ⁴⁰	cat
	inau_tiempoCompleto	Indica si ha sido inscripto durante el 2021 en un centro de INAU que trabaje en la modalidad de tiempo completo con menores de edad que se encuentran separados de su familia de origen (transitoria o permanentemente).	cat
	Fonasa	Indica si tiene cobertura por Fonasa ⁴¹	cat
	Afam	La familia es beneficiaria de asignaciones familiares ⁴²	cat
	tus	El alumno integra un hogar que percibe la Tarjeta Uruguay Social. ⁴³	cat
Datos de la persona a cargo	adulto_parentesco	Parentesco que tiene con la persona adulta que se encuentra a su cargo.	cat
	adulto_extranjero	Indica si esta persona adulta es uruguaya o extranjera. Esta variable fue incluida porque fue la que podría recabar de	cat

³⁸“El Programa de Acogimiento Familiar es un recurso de apoyo temporal, que tiene alcance a nivel nacional. Cuando la convivencia con la familia de origen del niño, niña o adolescente ya no es posible, se puede solicitar asistencia a INAU. Desde la Institución, se buscará una solución para asegurar su desarrollo en un nuevo entorno familiar, procurando estimular el respeto de su identidad y sus vínculos afectivos.” *Programa de Acogimiento Familiar*. (2018, enero 15). Gub.uy. <https://www.inau.gub.uy/familia/acogimiento-familia>

³⁹“Este sistema atiende a la infancia en situación de calle a través de diversas estrategias de intervención con el fin de contribuir a la promoción, protección y restitución de derechos de niños/as, adolescentes y sus familias en situación de alta vulneración de derechos. En el marco de este sistema, se realizan actividades de captación y acompañamiento en calle a través de equipos móviles, actividades diurnas de seguimiento y acompañamiento en distintos centros de referencia, y en caso que corresponda, el sistema cuenta también con atención en modalidad de tiempo completo.” *Sistema de Respuesta para Situación de Calle*. (2014, mayo 29). Gub.uy. https://www.inau.gub.uy/content_page/45-programas-asociados-a-situaciones-de-vulneracion-de-derechos/1044-sistema-de-respuesta-para-situacion-de-calle

⁴⁰“Los programas de Atención las Situaciones de Calle se dirigen a personas que hacen uso de centros (refugios) en los que se brinda alimento, alojamiento y abordaje multidisciplinario que colabore en la adquisición de hábitos de la vida cotidiana, cuidados de salud y recuperación de redes sociales. Existen 3 tipos de centros de atención: para mujeres con menores a cargo, para personas que requieren cuidados y para personas en situación de calle (propriadamente dicho). También se identifica y registra a las personas atendidas por equipos móviles del PASC.” *Atención integral*. (s/f). Ministerio de Desarrollo Social. <https://www.gub.uy/ministerio-desarrollo-social/comunicacion/publicaciones/atencion-integral>

⁴¹“El Sistema Nacional Integrado de Salud reglamenta el derecho a la protección de salud de todos los habitantes del país, a través de un Seguro Nacional de Salud (SNS), financiado por el Fondo Nacional de Salud (Fonasa), extendiendo obligatoriamente la cobertura médica de los trabajadores a su núcleo familiar a hijos a cargo menores de 18 años o mayores con discapacidad, propios así como los del cónyuge o concubino cuando no estén amparados por sus padres biológicos o adoptivos, o cónyuge o concubino que no posea por sí mismo la cobertura médica del Seguro Nacional de Salud (SNS).” Banco de Previsión Social. *Fondo Nacional de Salud (Fonasa)*. Banco de Previsión Social. <https://www.bps.gub.uy/10310/fondo-nacional-de-salud-fonasa>

⁴²“Es una prestación económica bimestral que BPS brinda a los hijos o menores a cargo de los trabajadores. Se otorga desde el momento de comprobado el embarazo y varía según los ingresos del hogar. Aumenta según la cantidad de hijos.” Banco de Previsión Social. *Asignación familiar*. Banco de Previsión Social. <https://www.bps.gub.uy/5470/asignacion-familiar.html>

⁴³“Es una transferencia monetaria que se otorga a aquellos hogares en situación de extrema vulnerabilidad socioeconómica. Su principal objetivo es asistir a los hogares que tienen mayores dificultades para acceder a un nivel de consumo básico de alimentos y artículos de primera necesidad.” *Tarjeta Uruguay Social*. (s/f). Ministerio de Desarrollo Social. <https://www.gub.uy/ministerio-desarrollo-social/politicas-y-gestion/programas/tarjeta-uruguay-social>

		mejor manera la condición de migrante de la familia, ya que este no es un dato que se encuentre sistematizado en los sistemas de la ANEP.	
	edadAdultoNacimiento	Edad que tenía la persona a cargo en el momento del nacimiento de este estudiante. (Se realiza el cálculo a partir de la fecha de nacimiento de ambos indicada en SIAS).	num
	adulto_trabajoFormal	Indica si la persona a cargo tiene un empleo formal (en blanco). Esto se desprende de si está registrada en BPS ⁴⁴ . También se recabó esta información del SIAS.	cat
	adulto_subsidioDesempleo	La persona a cargo está cobrando un subsidio por desempleo.	cat
	adulto_subsidioEnfermedad	Muestra si la persona a cargo cobra un subsidio por estar trabajando de forma activa pero por razones médicas se encuentra en el momento imposibilitada de trabajar.	cat
	adulto_subsidioMaPaternidad	La persona a cargo está cobrando un subsidio por maternidad o paternidad	cat
	adulto_asisteRefugio	Indica si la persona a cargo está asistiendo a centros de MIDES en los que se brinda alimento, alojamiento y abordaje interdisciplinario para la adquisición de hábitos de la vida cotidiana cuidados de salud y recuperación de redes sociales	cat
	adulto_jubilacion	Muestra si la persona a cargo cobra jubilación por haberse retirado de la actividad remunerada.	cat
	adulto_pensionInvalidez	La persona a cargo cobra mensualmente una pensión no contributiva por contar con dictamen del área Medicina Laboral del BPS que determina que tiene discapacidad común en situación de carencia de recursos para hacer frente a sus necesidades vitales o es una persona con discapacidad severa.	cat
	adulto_pensionVejez	La persona a cargo cobra mensualmente una prestación no contributiva por ser mayor de 70 años y encontrarse en una situación de vulnerabilidad socioeconómica.	
	adulto_pensionFallecimiento	La persona a cargo estaba vinculada con una persona jubilada, trabajadora, o amparada por subsidios de BPS y luego de su fallecimiento comenzó a cobrar una pensión brindada por BPS.	cat
Variable a predecir	Objetivo	Esta variable es la que se debe predecir. Indica si un estudiante repitió al final del año escolar 2021.	cat

Para generar el dataset se seleccionó a aquellos estudiantes que asistían a educación primaria entre 1ro y 6to año en 2021. Más allá de que en Uruguay la educación es obligatoria a partir de los 4 años, es a partir de 1er año de primaria que pueden tener experiencias de repetición.

⁴⁴“El Banco de Previsión Social un ente autónomo creado en el artículo 195 de la Constitución de la República Oriental del Uruguay del año 1967. Su cometido central es coordinar los servicios estatales de previsión social y organizar la seguridad social.” Banco de Previsión Social, B. (s/f-b). *Evolución histórica*. Banco de Previsión Social. <https://www.bps.gub.uy/11626/evolucion-historica.html>

Se quitaron estudiantes que asistieron entre 2019 y 2021 a escuelas de educación especial⁴⁵, ya que estos centros educativos tienen características distintas al resto. Actualmente asisten a esta oferta educativa 5.558 estudiantes.

Se removió también a quienes no tenían información histórica para analizar, probablemente porque asistieron previamente a centros de educación privada.

Nos quedamos entonces con un total de 229.201 estudiantes para analizar y 111 variables.

Queda pendiente ver cómo se podrían analizar a esta parte del estudiantado que quedó fuera del dataset.

2.2. Análisis descriptivo

Se realiza un análisis para el 2021 de aquellas variables que intuitivamente se considera que podrían llegar a tener incidencia en la repetición de un estudiante.

Por un lado, se distingue según el departamento donde se ubica el centro educativo y se observa una mayor repetición de la matrícula de primaria en Montevideo en relación a los del interior del país.

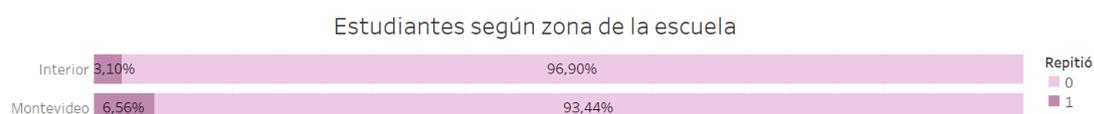


Figura 4. Estudiantes según zona de la escuela

Por otro lado, si se toma en cuenta el quintil socioeconómico en que se ubica la escuela, observamos como la repetición va disminuyendo a medida que aumenta el quintil.

A su vez, si se agrega la variable género (1= mujer, 0=hombre), podemos ver que sin importar el quintil, los hombres presentan mayores tasas de repetición.

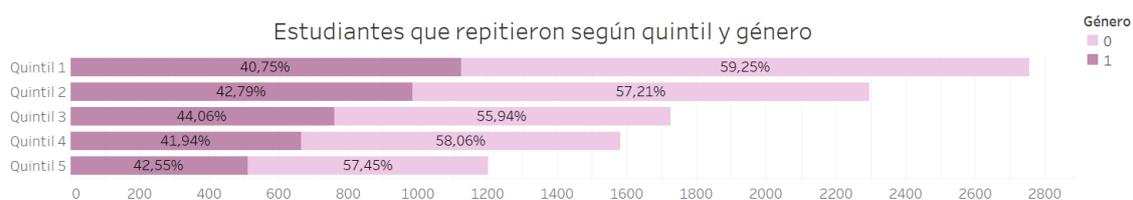


Figura 5. Estudiantes que repitieron según quintil y género

Por otro lado, si observamos la parte del estudiantado cuyas familias son beneficiarias de la Tarjeta Uruguay Social (TUS), que son familias en situaciones de extrema vulnerabilidad, se observa que quienes perciben esta prestación presentan una mayor repetición que quienes no la perciben.

⁴⁵Las escuelas especiales son “una red de escuelas, aulas, unidades de apoyo, proyectos, equipos, dispositivos, estrategias, recursos, modalidades de la Educación Inicial y Primaria, para favorecer los aprendizajes de niños, niñas y adolescentes con discapacidad, problemas para aprender y altas capacidades, desarrollada desde una perspectiva de derechos, en el marco de una educación inclusiva y de una Escuela con todos, para todos y cada uno”. Educación especial. (s/f). Edu.uy. <https://www.dgeip.edu.uy/educacion-especial-es/>

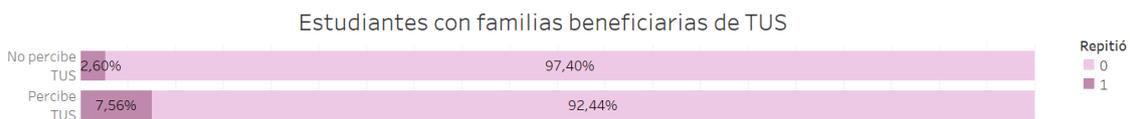


Figura 6. Estudiantes con familias beneficiarias de TUS

Se analiza también si influye en el desempeño de los estudiantes la situación laboral de la persona a su cargo. Según lo que se observa, quienes tienen una persona a cargo con trabajo formal tienen una menor repetición (3,08%) que quienes no tienen una persona a cargo con trabajo en blanco (5,42%).



Figura 7. Estudiantes según la situación laboral de la persona a cargo

Analizando la edad que tenía la persona a cargo al momento del nacimiento del estudiante, podemos observar que quienes tienen una mayor repetición son estudiantes con personas a cargo que tenían menos de 18 años cuando nació el alumno, o que tenían más de 65 años en ese momento.

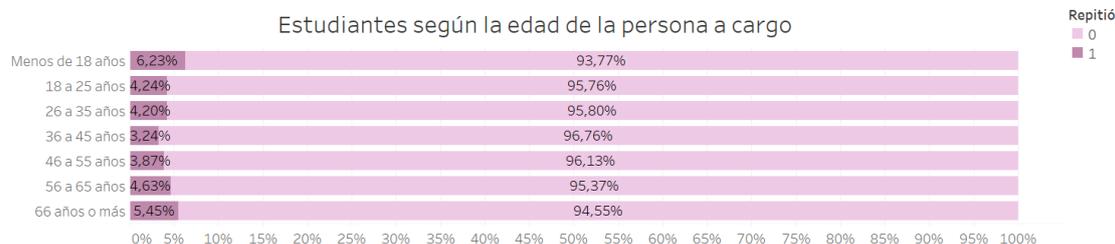


Figura 8. Estudiantes según la edad de la persona a cargo

3. Metodología

La aplicación de modelos de machine learning es algo muy novedoso para el sistema educativo uruguayo, por lo que no es menor tomar en cuenta las resistencias que esto puede llegar a tener en el plantel docente. Implicaría un trabajo profundo y minucioso con todos los actores involucrados, para compartir los beneficios y posibilidades que tendría la aplicación de este tipo de metodologías en la educación formal.

Siguiendo esta lógica, se decidió utilizar modelos basados en árboles. Estos modelos son flexibles y capaces de manejar variables de diversos tipos (cualitativas y cuantitativas), pero principalmente son más fáciles de explicar y de ser entendidos por un público sin conocimientos previos del área. Se puede explicar cuáles fueron las variables que tienen más incidencia en los resultados y pueden ser más amigables de entender para un público que no está habituado a trabajar con estas herramientas cuantitativas.

3.1. Ingeniería de atributos

Las variables categóricas del dataset presentan, en general, una cantidad muy alta de categorías posibles. A efecto de poder estimar los modelos que se proponen de manera efectiva se decidió realizar agrupaciones convenientes de las mismas. Para mayor detalle sobre los criterios utilizados referirse al apéndice (página 35) en donde se describen de forma breve los criterios empleados para agrupar.

Localidad

Hay 472 localidades, las cuales se relacionan cada una con un departamento (Uruguay tiene 19 departamentos). Se intentó aplicar Regression Trees para poder agrupar a las mismas pero no se llegó a buenos resultados, el modelo no predecía muy bien la variable objetivo a partir de las localidades, por lo que se optó por eliminar esta variable espacial y ver si en un futuro se puede complementar con más información, por ejemplo agrupando las localidades según su cantidad de habitantes.

Faltas

Para las variables asociadas con las faltas de un estudiante, que son faltas justificadas, faltas injustificadas y faltas totales (la suma de las dos anteriores), se hizo un análisis acerca de la distribución de las mismas en cada uno de los años 2021, 2020 y 2019. A partir de esto y tomando en cuenta el promedio de clases semanal y mensual, se agruparon las mismas.

En el apéndice se brinda más detalle acerca de este proceso.

3.2. Modelos a explorar

Para modelar la probabilidad de repetición en los estudiantes se utilizarán árboles y modelos ensamblados basados en árboles (Random Forest y Boosting) por los siguientes motivos: (1) Para capturar patrones no lineales entre las variables, (2) para poder entrenar modelos con relativamente pocos hiperparámetros (dado el tamaño de la base y las restricciones computacionales), (3) por la naturaleza mixta de la base de datos (variables numéricas y categóricas con muchas categorías) y finalmente (4) para evitar problemas de multicolinealidad y de tener que hacer supuestos fuertes en la distribución que siguen los datos.

A continuación se discuten técnicas de selección de variables que nos permiten hacer factible la estimación de los modelos antes mencionados renunciando a la menor cantidad posible de capacidad predictiva.

3.2.1. Feature selection

El data set original tiene 110 variables, de las cuales muchas son categóricas y tienen asociadas cientos o miles de categorías. En adición a los tratamientos que se comentaron en la sección 3.1 (ingeniería de atributos), también se propone utilizar un modelo logit regularizado mediante lasso con el fin de seleccionar variables relevantes para predecir cuáles son los alumnos van a repetir de curso atendiendo no solo a principios de parsimonia al entrenar los modelos, sino también para poder llevar adelante el fitting de los mismos dado los recursos computacionales de los que se dispone. Este modelo auxiliar, con el que se seleccionan features relevantes, se

fittea eligiendo el hiperparámetro de regularización por validación cruzada, y de esta manera de entre las 110 variables originales logramos descartar 80; es decir que los modelos de árbol ensamblados (Forest y Boosting) se van a fittear con las 30 variables más relevantes seleccionadas con Lasso.

3.2.2. Random Forest

El modelo Random Forest ensambla las opiniones de varios modelos de árbol promediando sus predicciones (o sometiendo a votación por mayoría las predicciones de los modelos en caso de problemas de clasificación multiclase). El error de generalización (o error fuera de la muestra de entrenamiento) de este modelo dependerá de dos parámetros sensibles: El tamaño (complejidad) de los árboles que ensamblamos y de un hiperparámetro que determina que cantidad de features tomamos en cuenta cada vez que se necesita estimar una 'regla' para cada árbol del bosque. En general estos modelos presentan un mejor desempeño que un árbol en solitario, ya que al promediar una cantidad grande de predictores se produce un efecto de regularización el modelo (menos varianza, mismo sesgo que un árbol en solitario). A cambio de este mayor poder predictivo, los ensambles Forest no se pueden interpretar gráficamente (como si ocurre con los árboles tradicionales), sin embargo sí es posible calcular métricas de importancia de cada variable lo que permite entender cuáles son los factores relevantes que a criterio del modelo determinan la probabilidad de que un estudiante repita de curso (más abajo hacemos uso de estas herramientas).

Se aplicó entonces Random Forest sobre el dataset con las variables más importantes y se eligieron los valores de los hiperparámetros mediante validación cruzada (se estimó la tasa de error out of bag para varias combinaciones de los hiperparámetros sensibles y nos quedamos con la combinación que minimiza dicho error). Se consideró entonces un ensamble que en cada corte sortee 20 de entre las 30 variables relevantes para la predicción y que cada árbol del bosque debe tener como mucho 100 nodos terminales.

A efectos de mostrar de manera preliminar la capacidad predictiva del modelo, elegimos como umbral de corte el valor 0.08 y utilizando out-of-bag clasificamos a un estudiante como repitente si su probabilidad estimada (fuera de la muestra de train) es mayor o igual a 0.08. En la Tabla 2 damos cuenta de la matriz de confusión del modelo estimado con el umbral antes mencionado. En la sección 4 se elaboran criterios específicos para elegir el umbral de corte de acuerdo con los objetivos planteados en este trabajo.

Prob>0.08	0	1	class.error
0	43422	566	0.11333333
1	1642	210	0.88666666

Tabla 2. Matriz de confusión para Random Forests

Como podría esperarse, con un umbral de corte relativamente elevado detectamos solo al 27% de los estudiantes que van a repetir, mientras que estaríamos tomando acciones de forma innecesaria sobre un 4% de la matrícula, que representan unos 7600 estudiantes.

Se observó que aquellas variables que tenían mayor importancia eran las siguientes;

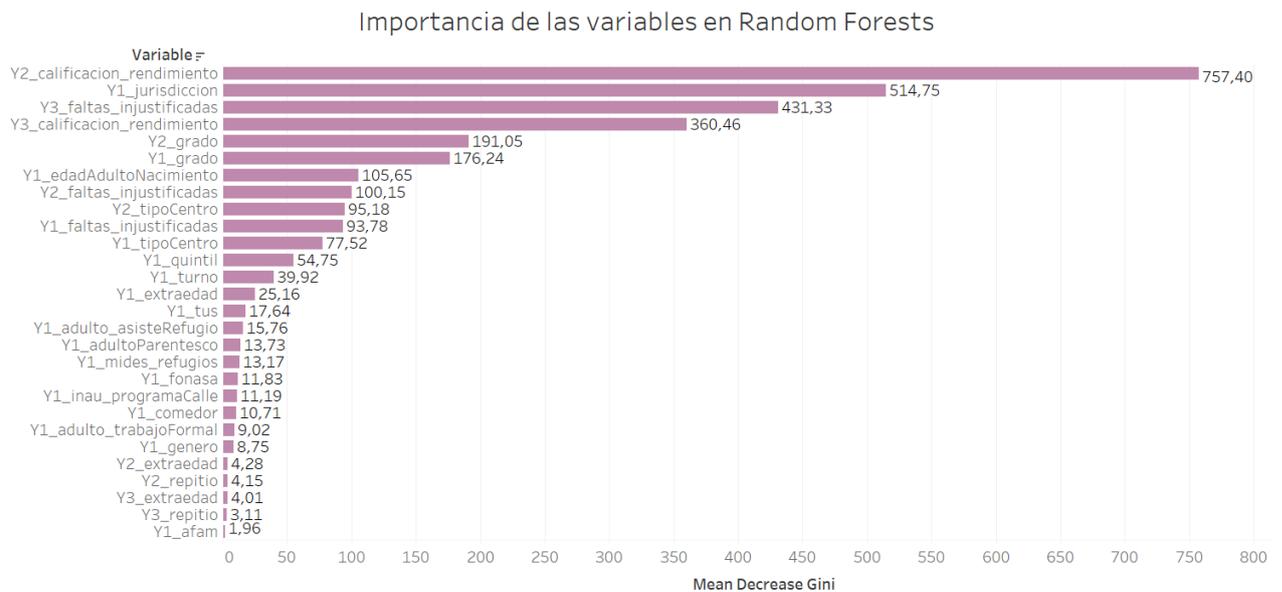


Figura 9. Importancia de las variables en Random Forest⁴⁶

Tal como se observó en el análisis previo de los datos, la jurisdicción en el que se encuentra el centro es de gran importancia sobre el desempeño del alumno, pero el hecho de que la familia perciba la Tarjeta Uruguay Social no se ubica entre las diez variables más importantes. A su vez, llama la atención también que el quintil del centro educativo no es de las variables que tienen una mayor incidencia.

3.2.3. Boosting machines

El modelo Forest construye el ensamble considerando re-muestras con reemplazamiento de la muestra original. En este sentido, cada árbol del ensamble Forest se fittea por separado (de manera independiente de los demás) con el fin de luego poder promediar la opinión (las predicciones) de cada uno de los árboles del ensamble. Si bien cada árbol puede en principio ser relativamente extenso (poco sesgo y mucha varianza) al promediar los modelos se produce un efecto de regularización (menor varianza). Por el contrario, y a diferencia de Forest, las Máquinas Boosting ensamblan una secuencia de modelos relativamente simples e interdependientes. Cada árbol dentro de Boosting tiene poca varianza pero alto sesgo, sin embargo al ensamblar modelos a lo largo de una secuencia interdependiente de modelos la varianza del ensamble permanece relativamente estable y logramos disminuir el sesgo.

Los árboles Boosting tienen varios parámetros sensibles; entre los que resultan más importantes de ajustar se encuentran:

1. El número de árboles B . Boosting puede hacer overfitting si B es muy largo, aunque este overfitting tiende a ser lento.
2. El parámetro de encogimiento λ , un número pequeño y positivo. Este controla el ratio al cual Boosting aprende. Los valores típicos son 0,01 o 0,001. λ muy

⁴⁶ Y1 refiere a las variables para 2021, Y2 a las de 2020 e Y3 a las variables del año 2019

pequeños pueden requerir usar un valor muy grande de B para poder alcanzar buena performance.

3. El número de splits d en cada árbol, lo que controla la complejidad del conjunto de la muestra. Normalmente $d = 1$ funciona bien, en ese caso cada árbol es un stump, que consiste en un solo Split.⁴⁷

Para fitear el modelo boosting, trabajamos con el mismo data set que utilizamos para el modelo anterior, y también se eligieron los valores de los hiperparámetros sensibles del modelo utilizando técnicas de validación cruzada. A continuación se observa que manteniendo como punto de corte 0,08 con 620 árboles y un shrinkage de 0,01 se tiene un error rate de 7,91% pero en relación a la captación de alumnos 1 esta es de 86,93%;

Prob>0.08	0	1	class.error
0	41792	820	0.1307311
1	2806	422	0.8692689

Tabla 3. Matriz de confusión para Boosting

Se podría entonces con este modelo y punto de corte, captar al 34% de los estudiantes que van a repetir, y se tomarían acciones sobre un 6% de estudiantes que no van a repetir. Estos representarían unos 13200 estudiantes.

A continuación se presentan las variables que observa son las más importantes;

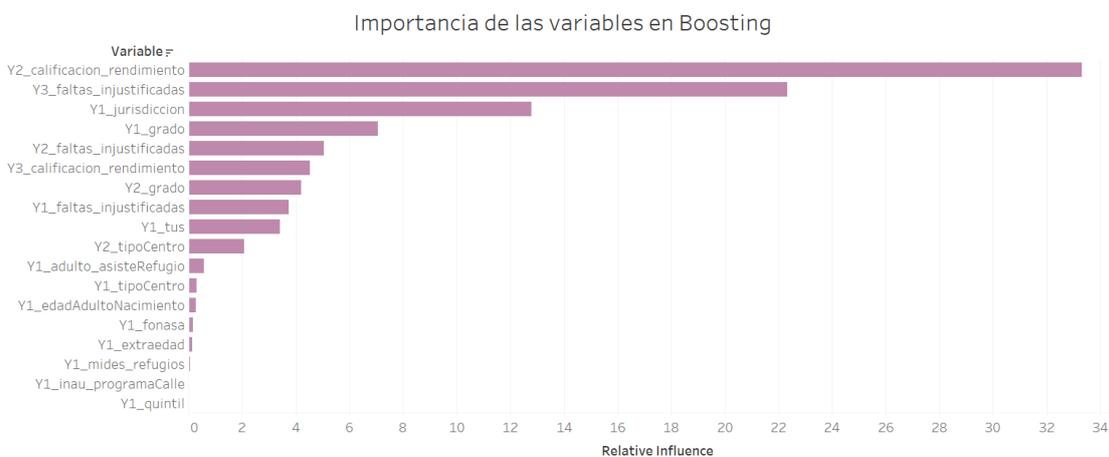


Figura 10. Importancia de las variables en Boosting Machines

Observamos que la calificación en rendimiento es la variable más importante en los dos modelos comparados. También lo son las faltas injustificadas, que en Random Forests ocupan el tercer lugar en importancia y en Boosting Machines ocupan el segundo lugar.

La jurisdicción donde se encuentra el centro educativo también se repite como variable de importancia. En Random Forests ocupa el segundo lugar y en Boosting Machines, el tercero.

Otras variables de interés que se repiten en ambos modelos son el grado en que se ubica el alumno, y si este repitió.

⁴⁷James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021).

Por otro lado, en relación las diferencias, se observa que Random Forest coloca en un lugar de importancia a la edad del adulto al momento del nacimiento y al tipo de centro, mientras que Boosting no, pero suma el hecho de si la familia del alumno percibe TUS o no, y si el adulto asiste a algún refugio.

A continuación se evaluará entonces cuál es el punto de corte óptimo sobre el cual trabajar en cada uno de los modelos para poner en línea los outputs de ambos modelos con el objetivo del trabajo.

4. Resultados de la modelización y estrategias a aplicar

4.1. Resultados de la modelización

En el punto 4.2 se discutirá acerca de qué tipo de acciones se podrían llegar a tomar sobre el conjunto de estudiantes con mayor riesgo de repetición, para así minimizarlos. Las acciones que se lleven a cabo hacia este grupo particular serán distintas según las características del centro educativo al que asista, y según las particularidades de sus familias y su contexto.

En esta sección nos concentraremos en discutir cómo elegir el punto de corte (umbral) para clasificar a un alumno como repitente, tomando en cuenta que resulta complejo cuantificar cuál sería el costo de tomar acciones sobre un estudiante que no es vulnerable, en relación a dejar de tomar acciones sobre un alumno que sí lo es, pero dadas las consecuencias que se detallaron en la primera parte de este trabajo que puede tener la repetición sobre la vida de un niño, niña o adolescente, se considera que un falso positivo (repitente) predicho por el modelo es mucho menos costoso que un falso negativo (no repitente).

A efectos de poder comparar entre los modelos, en la figura 11 se muestran, para distintos umbrales y modelos, los porcentajes de verdaderos positivos definidos como:

$\% \text{ Verdaderos positivos} = \frac{\text{estudiantes repitentes correctamente identificados por el modelo}}{\text{total estudiantes repitentes}}$

Y también se muestran los porcentajes de falsos negativos, definidos como:

$\% \text{ Falsos Positivos} = \frac{\text{estudiantes que los modelos reconocen como repitentes y no repiten de curso}}{\text{total alumnos no repitentes}}$

En base a esta información podemos observar que ambos modelos no solo ordenan la importancia de las variables de forma diferente, sino que existen diferencias significativas respecto de que predice mejor cada modelo. Resulta interesante notar que para cualquier umbral de corte que elegimos, el modelo Boosting captura sistemáticamente un porcentaje de verdaderos positivos más grandes que Random Forest, mientras que Random Forest presenta sistemáticamente siempre un menor porcentaje de falsos negativos.

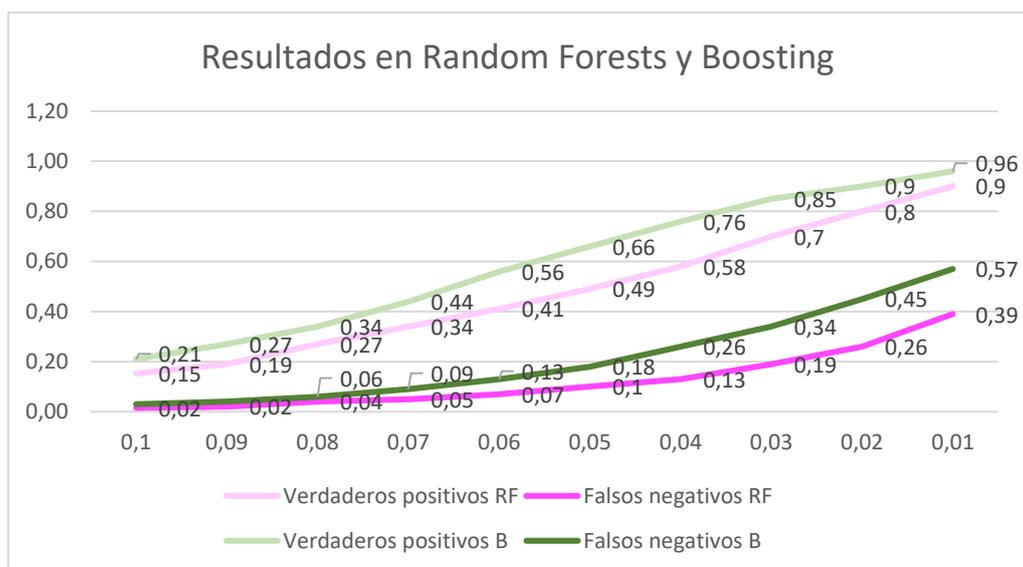


Figura 11. Resultados en Random Forests y Boosting

A efectos de dirimir que modelo y umbral utilizar en la práctica a la hora de embeber las predicciones del modelo en el marco de un programa de apoyo a los estudiantes en riesgo de repetir de curso, se plantean dos preguntas relevantes:

1. “¿Cuál es el umbral de corte que necesita cada modelo para capturar al menos al 90% de los alumnos que van a repetir de grado?”
2. “¿Entre ambos modelos y para el umbral que responde a la pregunta (1), cuál de los dos modelos garantiza que se produce el menor porcentaje de falsos negativos?”

Responder a la pregunta (1), es decir, la elección de un umbral que garantice cubrir al menos al 90% de estudiantes repitentes se puede considerar un objetivo de política educativa, mientras que la selección del modelo que cumpliendo con este objetivo también minimice la tasa de falsos negativos (pregunta (2)) surge simplemente de la necesidad de implementar criterios de eficiencia en el trazado de las políticas públicas (se pretende minimizar el costo de implementar las ayudas que se discuten en la próxima sección).

Mirando la gráfica 11, resulta entonces claro que al responder las preguntas (1) y (2), deberíamos considerar el modelo Random Forest y un umbral de corte de (aproximadamente) 0.01; esto garantiza alcanzar una tasa de acierto del 90% y una tasa de falsos negativa del orden del 39%. Esto quiere decir que se aplicarían políticas de apoyo innecesarias sobre (aproximadamente) unos 85 mil alumnos, mientras que si utilizamos el modelo Boosting, el apoyo innecesario se daría sobre (aproximadamente) un total de 130 mil alumnos.

4.2. Estrategias a aplicar

Una vez identificados los sujetos del estudiantado en riesgo de repitencia (utilizando el modelo y los criterios definidos en la sección anterior) se plantea tomar acciones sobre este conjunto para mejorar sus trayectorias educativas. De esta manera, no solo prevenimos futuras instancias de repetición, sino también una mejor dinámica en la eficiencia administrativa de los recursos del organismo en los próximos años, ya que el hecho de que un estudiante repita genera un

aumento del costo por alumno dentro del sistema, asumiendo que logrará culminar la trayectoria escolar.

En esta tesis planteamos que en una primera etapa se ensayen distintas acciones, descritas a continuación, sobre grupos pequeños de en riesgo de repetición de forma aleatoria para poder medir la eficacia e impacto de cada estrategia de acompañamiento escolar (a cada grupo se le aplica un tratamiento diferente y luego se estima que tratamientos tienen mayor efectividad y cuales conviene repensar o descartar). De esta manera se podrían ir evaluando las actividades llevada a cabo y observar si están teniendo resultados o no, al mismo tiempo de medir cuáles de los tratamientos fueron más eficaces en su cometido de prevenir la repetición escolar, y así poder implementarlos a gran escala, generando entonces una política pública basada en evidencia científica.

Entre las acciones de acompañamiento se pensaron en las que se describen a continuación, las cuales se ordenan desde las que se cree serían más viables de incorporar en el mediano plazo, a las menos viables.

4.2.1 Aprovechamiento de plataformas disponibles

A nivel práctico, es necesario que la información acerca quienes son estos estudiantes más vulnerables esté disponible de manera sencilla para todas las personas que forman parte centro educativo. Actualmente en primaria se trabaja con el sistema GURI, una aplicación que el docente puede tener en su computadora o teléfono para que registre asistencias y comentarios sobre el alumno. Por la dinámica de trabajo, esta aplicación es abierta por los maestros todos los días, así que tendría sentido disponibilizar ahí la información generada a partir de este estudio.

Por otro lado, los docentes trabajan también con plataformas educativas creadas por el Plan Ceibal.

El Plan Ceibal “es el centro de innovación educativa con tecnologías digitales del Estado uruguayo, al servicio de las políticas públicas educativas. Promueve la integración de la tecnología a la educación con el fin de mejorar los aprendizajes e impulsar procesos de innovación, inclusión y crecimiento personal.”⁴⁸ Hay actualmente tres plataformas a través de las cuales los docentes pueden trabajar con sus alumnos; CREA⁴⁹, Matific⁵⁰ y la Plataforma de Lengua⁵¹. En línea con esto, existen estudios que muestran que la mayor utilización de software en el aprendizaje tiene un impacto positivo sobre los alumnos.⁵²

Más allá del trabajo particular que haga el docente con cada una de estas plataformas, no existen actualmente lineamientos específicos acerca de cómo deben ser utilizadas, sino que queda a consideración del plantel docente. Es a partir de esto que se propone que quienes tengan una mayor probabilidad de repetición tengan más tareas y más ayuda en la utilización de estas plataformas, reforzando así el aprendizaje en las áreas donde tengan mayor dificultad, ayudándolos de esta manera en la mejora de sus resultados académicos.

⁴⁸Plan ceibal - sobre nosotros. (s/f). Edu.Uy. <https://www.ceibal.edu.uy/es/institucional>

⁴⁹Plan ceibal - Crea. (s/f). Edu.Uy. <https://www.ceibal.edu.uy/crea>

⁵⁰Matific. (s/f). Edu.Uy. <https://www.ceibal.edu.uy/matific>

⁵¹Plataforma de Lengua. (s/f). Edu.uy. <https://www.ceibal.edu.uy/plataformadelengua>

⁵²SkillsBank una herramienta para explorar programas efectivos de desarrollo de habilidades. (s/f). iadb.org. [https://skillsbank.iadb.org/es/policy-challenge/learning+in+primary/guided+technology+\(with+extra+time\)](https://skillsbank.iadb.org/es/policy-challenge/learning+in+primary/guided+technology+(with+extra+time))

A su vez, Plan Ceibal ha creado la Biblioteca País, la cual ofrece en forma gratuita y para cualquier dispositivo, contenidos para la comunidad educativa y para la población en general. Busca una democratización del acceso a la lectura y a la cultura al ofrecer libros de diversa índole.⁵³

Existen diversos estudios acerca de cómo el hábito lector incide en la repetición, generando una baja entre aquellos estudiantes que leen y una suba entre aquellos que no leen, ampliándose también esta incidencia hacia las familias de los estudiantes, más allá de cuál sea la situación socioeconómica de la misma.⁵⁴

En esta línea, se pueden realizar talleres con las familias para la estimulación de la lecto-escritura, los cuales pueden también ser ejecutados con especialistas invitados fuera de la institución.⁵⁵

4.2.2. Trabajo en conjunto con las comunidades

Finalmente, es importante destacar el hecho de que entre las comunidades educativas podría llegar a haber ciertas resistencias al utilizar modelos de aprendizaje automático aplicados a la población escolar. Algunos de los argumentos hablan de que esto podría generar un cierto estigma sobre el estudiante y que pueda crear un impacto negativo en su motivación para la continuidad de sus trayectorias académicas.

Es en esta línea es que se propone un trabajo en conjunto con las comunidades que se verán involucradas en este tema. Una vez que se tengan los resultados se podría trabajar con las inspecciones y el equipo de protección de trayectorias, para concurrir a los centros educativos y hablar con docentes y equipos directivos.

A partir de esto se podrían generar instancias de intercambio acerca de qué son este tipo de modelos, que resultados tienen y qué beneficios pueden llegar a tener, siempre teniendo en cuenta que el objetivo es que esta sea solamente una herramienta más para apoyarlos a poner un mejor foco en el estudiante y así prevenir su repetición y posterior desvinculación del sistema educativo.

4.2.3 Incorporación del equipo del Sistema de Protección de Trayectorias

En el 2015 el CODICEN de la ANEP creó el Sistema de Protección de Trayectorias Educativas (SPTe), entendiéndose que este sistema permitiría realizar un mejor seguimiento de los estudiantes, sobre todo de quienes tienen un inminente riesgo de abandono escolar. El sistema considera necesario proteger y acompañar la educación obligatoria, y está enfocado a la educación media básica.

A partir de esto se crearon las Unidades Coordinadoras Departamentales de Integración Educativa (UCDIE), con el fin de proyectar la educación en cada uno de los territorios. Estas Unidades se incorporaron a aquellos espacios organizados en los territorios, a nivel local, departamental o regional y trabajan junto a los jóvenes de educación media básica del territorio en el que se encuentran, para acompañar sus trayectorias educativas.

⁵³Biblioteca País - Plan Ceibal. (s/f). Edu.uy. <https://www.ceibal.edu.uy/biblioteca>

⁵⁴Cabrera, L. (2019)

⁵⁵Campos, N., Méndez, E., Presa, T., & Habiaga, V. (2018)

A pesar de que estas Unidades trabajan con educación media, tienen amplia experiencia en lo relacionado a la articulación entre las familias y los centros educativos, y en el seguimiento personalizado de estudiantes. Se propone entonces que para el grupo de estudiantes identificados como vulnerables, se articule con estas UCDEs para que den apoyo a los centros educativos en la atención a estos estudiantes.

4.2.4 Evaluaciones y seguimiento

Realización de evaluaciones diagnósticas de forma individual, así como seguimientos de los aprendizajes más de cerca de lo normal. Se pueden realizar evaluaciones diarias y seguimientos a través de registros acumulativos.

Se propone que los docentes realicen evaluaciones (además de las realizadas normalmente) a estudiantes que están en riesgo, para así realizar un monitoreo más de cerca. Además, si se sistematiza toda esta información, se pueden tomar mejores acciones a lo largo del año escolar.⁵⁶

4.2.5 Tutorías individuales

Existe actualmente amplio consenso acerca de que las tutorías individuales generan un impacto sobre estudiantes con diversos problemas. Se afirma que estas “inyectan verdadero dinamismo al proceso de aprendizaje”⁵⁷ por lo que se propone que estudiantes que tengan alta probabilidad de repetición, tengan durante el año escolar un seguimiento individualizado que se manifieste en forma de tutorías, para así acompañar su trayectoria y disminuir esta posibilidad de repetición.

A su vez, el modelo generado a partir de este trabajo permitiría identificar cuáles son los centros que tienen una mayor cantidad de alumnos vulnerables, pudiéndose entonces tomar acciones con foco en el centro y mejorar el entorno escolar trabajando sobre el sentido de pertenencia.

Se afirma muchas veces que las escuelas pueden ser las generadoras de problemas en relación a la desvinculación escolar, pero que también pueden ser la solución. Se debería trabajar entonces en generar escuelas más inclusivas y que se adapten a las necesidades diversas de cada uno de los estudiantes. Algunas de las acciones recomendadas en esta línea tienen que ver con planes de capacitación dirigidos hacia el plantel docente y el agrupamiento de estudiantes que enfrentan desafíos similares⁵⁸

4.2.6 Participación de la familia

Es necesario apoyar a las comunidades educativas en la generación de espacios para actividades e iniciativas que, poniendo al estudiante en el centro, permitan posicionar a la escuela como un espacio de intercambio y colaboración con sus familias.

⁵⁶Campos, N., Méndez, E., Presa, T., & Habiaga, V. (2018)

⁵⁷Messina, J. (2018)

⁵⁸Josephson, K., Francis, R., & Jayaram, S. (2018)

Hay amplia evidencia acerca de cómo la participación de las familias resulta crucial para un buen desarrollo del estudiante. Las actividades a realizar pueden ser más generales y preventivas, por ejemplo reuniones entre maestros y familias, o más específicas, por ejemplo si hay determinadas familias con un menor nivel de participación, realizar comunicaciones particulares con estas o hasta visitas en el hogar.

En línea con esto, es importante brindar lineamientos y orientaciones hacia las escuelas, pero no generar una sobre intervención dentro de las mismas, no se recomienda una sobrecarga de los actores dentro del centro para coordinar actividades, sino que lo ideal sería que las escuelas puedan trabajar de forma autónoma y apoyándose en sus conocimientos sobre las familias que asisten al centro para así brindar a las familias mejores y más eficaces respuestas.⁵⁹

4.2.7 Desestructuración del aula y enseñanza a través de lo lúdico

Otras propuestas se relacionan con la utilización de diversos espacios dentro y fuera del centro educativo, ya sea recorridos por la ciudad, salidas didácticas, o visitas a huertas. Se puede hacer énfasis en aquellos grupos donde hay una mayor proporción de estudiantes en riesgo y trabajar el aprendizaje basado en vivencias y la resolución de problemas reales que se pueden llegar a dar en las salidas didácticas.

También en lo que tiene que ver con desestructuración del aula, se puede trabajar dentro del mismo salón de clases pero en diversos espacios y fundamentalmente en grupos. Se pueden utilizar espacios más desestructurados como por ejemplo el patio de recreo. Por otro lado, se destaca los aspectos positivos del juego en este tipo de actividades.⁶⁰

4.2.8 Adaptaciones curriculares

Un tipo de estrategias de intervención que se identifica como posible es el de las adaptaciones curriculares. Esto incluye identificar a un grupo de estudiantes que tienen ciertas dificultades, (en este caso serían aquellos que el modelo identifica como más vulnerables) y pensar propuestas alternativas para ellos.

Otra opción en la misma línea es que un docente realice la misma propuesta para el grupo en general, y luego en la fase de trabajo la pueda adaptar de forma individual o en pequeños grupos.

Este tipo de propuestas están enfocadas en atender las diversidades dentro de las aulas así como a las particularidades de cada uno de los alumnos.⁶¹

4.2.9 Educación emocional

Otro tipo de acciones a tomar en relación a los estudiantes, que a menudo son olvidadas, tiene que ver con la inteligencia emocional de los estudiantes.

⁵⁹Josephson, K., Francis, R., & Jayaram, S. (2018)

⁶⁰Campos, N., Méndez, E., Presa, T., & Habiaga, V. (2018)

⁶¹Campos, N., Méndez, E., Presa, T., & Habiaga, V. (2018)

Existe un amplio consenso respecto a que una mayor inteligencia emocional repercute en un mejor rendimiento académico. Además, se observa que “cuando la enseñanza incorpora las emociones, los resultados son muy superiores a aquellos que se obtienen de aquel proceso impersonal y desprovisto de todo lo afectivo”⁶². Dentro de las habilidades que integran la inteligencia emocional se incluyen las habilidades del conocimiento de sí mismos, el auto control, la motivación, la empatía y las habilidades sociales.⁶³

Además, otros beneficios que se observan de tener un alto nivel de inteligencia emocional se vinculan con una mejoría en el bienestar psicológico, el cual influye sobre la salud mental de niñas y niños en edad escolar, adquiriendo como consecuencia, un equilibrio emocional potenciador del rendimiento académico. Esto genera una mejora no solo del rendimiento de estos estudiantes sino también del clima en las aulas.

A su vez, el estudio más a fondo de estas temáticas ha permitido encontrar ciertas causas que influyen sobre los procesos de aprendizaje y así trabajar en evitar por ejemplo los conflictos o el bajo rendimiento, al mismo tiempo que ha podido vincular a los procesos de aprendizaje con estados emocionales negativos⁶⁴

Una de las variables que influye de manera más relevante sobre el proceso de aprendizaje es la autoestima. Se ha comprobado que es un factor determinante en el rendimiento académico y social óptimo de los estudiantes. Se observa que el tener una mayor confianza en las propias capacidades, una saludable autoestima académica, al mismo tiempo que unas altas expectativas de autoeficacia generan una mayor valoración de las tareas y hacen que los estudiantes se sientan responsables de los objetivos de aprendizaje, mejorando así tanto su rendimiento como su estado emocional.⁶⁵

Por otro lado, el tener satisfactorias competencias emocionales también fomenta comportamientos positivos tales como la regulación de las emociones y una actitud positiva. Esto genera al mismo tiempo una reducción de los comportamientos contraproducentes o distorsionadores como el estrés o la ansiedad ante las evaluaciones o los conflictos. En concreto, se han observado mejoras en la conducta prosocial y una disminución de los comportamientos negativos y disruptivos, así como también una mejora en el bienestar del alumno”⁶⁶

Queda claro el hecho de que a mayor autorregulación emocional, mayor rendimiento. Por ejemplo, si no hay una autorregulación adecuada, no es posible para el alumno retener la información que el docente presenta en clase, generándole una dificultad para controlar su conducta en el aula. Además este descontrol puede incluir comportamientos agresivos y antisociales. Por otro lado, la falta de regulación emocional también influye sobre la recuperación de información y fomenta la ansiedad, afectando además sobre el rendimiento en las pruebas.⁶⁷

A raíz de esto, resta responder cómo pueden hacer los docentes para crear aprendizajes que sean emocionalmente significativos. Una estrategia consiste en ver cómo pueden influir sobre la regulación de la experiencia emocional de sus estudiantes. Por ejemplo, “puede observarse si el docente tiene un rol importante como guía de la respuesta emocional del estudiante: cada una de las acciones de este último podrían acompañarse de una conducta emocional, que podría

⁶²Acosta, F. P., y Clavero, F. H. (2017)

⁶³Acosta, F. P., y Clavero, F. H. (2017)

⁶⁴Morente, A. R., Guiu, G. F., Castells, R. R., & Escoda, N. P. (2017)

⁶⁵Morente, A. R., Guiu, G. F., Castells, R. R., & Escoda, N. P. (2017)

⁶⁶Morente, A. R., Guiu, G. F., Castells, R. R., & Escoda, N. P. (2017)

⁶⁷Gargurevich, R. (2007)

regularse de acuerdo a lo sugerido por el profesor. Por ejemplo, un educador puede seleccionar qué película será vista en el aula (selección de la situación), puede generar estímulos, puede crear argumentos para influir y modificar la evaluación del alumno (modificación de la evaluación) y, finalmente, puede suprimir sus respuestas emocionales. Es decir, el maestro puede ser guía y gestor del cambio del proceso emocional.”⁶⁸

En la Dirección General de Educación Inicial y Primaria se encuentra en funcionamiento el programa de Escuelas Disfrutables, el cual está integrado por equipos profesionales de psicología y trabajo social, que tienen como cometido realizar intervenciones interdisciplinarias en los centros educativos de ANEP.⁶⁹

Se propone entonces un trabajo articulado entre docentes y equipos de Escuelas Disfrutables para brindar atención a aquellos estudiantes identificados, y poder trabajar en conjunto para una mejora en sus habilidades socioemocionales.

Se ha observado que el trabajo en duplas pedagógicas y la planificación colaborativa entre docentes, equipos directivos y especialistas son valorados muy positivamente como estrategia preventiva de la repetición, así como se observa una influencia positiva sobre las prácticas pedagógicas a nivel colectivo.⁷⁰

Conclusiones

Tomando como base las altas tasas de repetición escolar del país, así como la elevada correlación que existe entre la deserción del sistema educativo formal y las experiencias de repetición en el ciclo de educación primaria, en este trabajo nos propusimos en primer término explorar diferentes modelos de aprendizaje automático para poder detectar (predecir) qué alumnos se encuentran en riesgo de repetir de curso durante su ciclo de educación primaria. Para ello, se utilizó una base de datos con más de 200 mil observaciones y 110 variables, que corresponde a casi todo el sistema de educación primaria pública del país. Con estos datos se llevaron a cabo análisis descriptivos del estado de situación del sistema educativo y se entrenaron dos modelos de aprendizaje automático que luego comparamos y calibramos para que a la hora de predecir cumplan con objetivos de política educativa y con criterios de eficiencia.

Finalmente, se discutió un conjunto de posibles cursos de acción a considerar, para tratar a los estudiantes en riesgo de repetir el año y se planteó la necesidad de experimentar en una prueba piloto cómo funcionarían estas estrategias para luego implementar sólo las más efectivas a nivel nacional. De esta forma una extensión natural del trabajo consiste en hacer un seguimiento sobre la implementación de estas políticas y profundizar en mejorar las mismas y/o idear nuevas estrategias de intervención que pudieran ser más efectivas.

Se considera también importante destacar las posibilidades de mejoras que podrían llevarse a cabo sobre los modelos entrenados. Particularmente, una de ellas sería realizar un mejor tratamiento sobre la variable localidad, al mismo tiempo que se destaca que no tener el dato de calificaciones para el año 2020 pudo haber impactado sobre la capacidad predictiva de los modelos. Es por esto que a futuro se plantea la necesidad de, a medida que pasen los años y se

⁶⁸Gargurevich, R. (2007)

⁶⁹Escuelas Disfrutables. (s/f). Edu.uy. <https://www.dgeip.edu.uy/programas/ped/>

⁷⁰Campos, N., Méndez, E., Presa, T., & Habiaga, V. (2018)

tengan más años lectivos con datos de calificación, se puedan imputar los datos faltantes en 2020 y mejorar así los resultados asociados a los modelos analizados.

Bibliografía

- Acosta, F. P., y Clavero, F. H. (2017). *LA INFLUENCIA DE LAS EMOCIONES SOBRE EL RENDIMIENTO ACADÉMICO*. Ciencias Psicológicas.
- Aristimuño, A., Giorello, M., y Lasida, J. (2020). *Ecosistema Nacional de I+D+i en Educación. Serie Working Papers SUMMA. N°2. SUMMA*
- Aristimuño, A. (2009). *El abandono de los estudios del nivel medio en Uruguay: un problema complejo y persistente*. REICE. Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación.
- Aristimuño, A., y Parodi, J. P. (2017). *Un Caso Real de Combate al Fracaso en la Educación Pública: Una Cuestión de Acompañamiento, Liderazgo y Cultura Organizacional*. REICE. Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación.
- Arroyo Resino, D., Constante Amores, I. A. y Asensio Muñoz, I. (2019). La repetición de curso a debate: un estudio empírico a partir de PISA 2015. *Educación XX1*, 22(2), 69-92, doi: 10.5944/educXX1.22479
- Bogliaccini, J., y Rodríguez, F. (2015). *Education system institutions and educational inequalities in Uruguay*. CEPAL Review 116.
- Bramer, M. (2007). *Principles of Data Mining*. Springer.
- Breiman, L. (2001). *Random Forests*. Kluwer Academic Publishers.
- Brunner, J. J. (2000). *Educación: escenarios de futuro. Nuevas tecnologías y sociedad de la información*. PREAL.
- Cabrera, L. (2019). *Políticas educativas preventivas de la repetición de curso en la enseñanza obligatoria en España*. Multidisciplinary Journal of Educational Research.
- Campos, N., Méndez, E., Presa, T., y Habiaga, V. (2018). *INFORME FINAL PROYECTOS I+D. Prácticas Pedagógicas Alternativas a la Repetición Escolar en primer año de la Jurisdicción Canelones Centro del CEIP*. Agencia Nacional de Investigación e Innovación.
- Gargurevich, R. (2007). *La autoregulación de la emoción y el rendimiento académico en el aula: El rol del docente*. Docencia Universitaria. Año 3, N°2
- INEEd. (2019). *Informe sobre el estado de la educación en Uruguay 2017-2018*. INEEEd.
- INEEd. (2021). *Informe sobre el estado de la educación en Uruguay 2019-2020*. INEEEd.
- INEEd. (2019). *La repetición en el sistema educativo uruguayo. Una mirada a la luz de los datos de la Encuesta Nacional de Adolescencia y Juventud (2008-2013)*. INEEEd.
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). *An Introduction to Statistical Learning*.
- Josephson, K., Francis, R., & Jayaram, S. (2018). *Políticas para promover la culminación de la educación media en América Latina y el Caribe. Lecciones desde México y Chile*. CAF; Banco de Desarrollo de América Latina.

Kučak, D., Juričić, V., & Đambić, G. (2018). *Machine learning in education - a survey of current research trends*. DAAAM International.

Messina, J. (2018, febrero 23). *Es hora de mejorar la educación primaria: ¿cómo lo hacemos?* Ideas que Cuentan. <https://blogs.iadb.org/ideas-que-cuentan/es/es-hora-de-mejorar-la-educacion-primaria-como-lo-hacemos/>

Morente, A. R., Guiu, G. F., Castells, R. R., & Escoda, N. P. (2017). *Análisis de la relación entre competencias emocionales, autoestima, clima de aula, rendimiento académico y nivel de bienestar en educación primaria*. REOP. Vol. 28, n°1.

Rivas, A. (2019). *¿Quién controla el futuro de la educación?* Siglo XXI Editores Argentina.

Román, M. (2013). *Factores asociados al abandono y la deserción escolar en América Latina: una mirada en conjunto*. REICE. Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación, vol.11

UNICEF. (2012). *Proteger las trayectorias escolares. Reflexión sobre la utilización de la repetición de grado*. UNICEF Uruguay.

Varela, J. P. (1874). *La educación del pueblo. Tomo I*. Sociedad de Amigos de la Educación Popular.

Apéndice A. Información adicional de variables

Jurisdicciones

A continuación se detallan las jurisdicciones según las cuales se ubican a los centros educativos.

Jurisdicción	Departamento
Montevideo - Este	Montevideo
Montevideo - Centro	Montevideo
Montevideo - Oeste	Montevideo
Artigas	Artigas
Canelones - Este	Canelones
Canelones - Pando	Canelones
Canelones - Costa	Canelones
Cerro Largo	Cerro Largo
Colonia	Colonia
Durazno	Durazno
Flores	Flores
Florida	Florida
Lavalleja	Lavalleja
Maldonado	Maldonado
Paysandú	Paysandú
Río Negro	Río Negro
Rivera	Rivera
Rocha	Rocha
Salto	Salto
San José	San José
Soriano	Soriano
Tacuarembó	Tacuarembó
Treinta y Tres	Treinta y Tres

Tabla A 1. Jurisdicciones

Tipos de centros educativos

Los centros educativos se dividen entre los siguientes tipos;

- Unidocente: Son los centros rurales que tienen un solo docente que dicta clases para los distintos grados de educación primaria
- Rural: Escuelas rurales que dictan clases durante 5 horas.
- Tiempo extendido: Son centros educativos donde se amplía el tiempo pedagógico. Durante la mañana se realizan talleres de educación física, danza, lenguas extranjeras y otras actividades. En la tarde los estudiantes tienen clases en su carga horaria habitual.
- Tiempo completo: Se imparte clases en el centro educativo durante 7 horas y media
- Aprender: Son centros ubicados en contextos de vulnerabilidad socioeconómica donde se imparte clases en la mañana o en la tarde. Se les brinda a los alumnos almuerzo y

desayuno o merienda y los docentes tienen instancias de coordinación junto a la figura del maestro comunitario para así atender mejor a las particularidades de los estudiantes

- De práctica: Son escuelas que buscan contribuir a la formación de los estudiantes de magisterio
- Común: Centros educativos donde los estudiantes tienen clases en la mañana o en la tarde

Quintiles

Con el objetivo de identificar a los centros con mayor vulnerabilidad y así priorizar áreas en los territorios donde realizar intervenciones específicas, se construye un índice de vulnerabilidad socioeducativa que se actualiza cada cinco años.

Para la construcción del mismo, la información que se toma en cuenta es; el porcentaje de estudiantes del centro educativo cuyos hogares perciben asignaciones familiares y el porcentaje de estudiantes del centro que son beneficiarios de la Tarjeta Uruguay Social, así como la localidad donde se encuentra el mismo.

Apéndice B. Ingeniería de atributos

Edad del adulto al momento del nacimiento

La decisión de calcular la variable de la edad del adulto a cargo en el momento del nacimiento, surge a raíz de la posibilidad de recabar información acerca de si el estudiante nació en circunstancias de mayor vulnerabilidad, ya sea a raíz de un embarazo adolescente, o por ejemplo, el no tener padres que hayan podido estar a su cargo y por eso están bajo el cuidado de una persona de mayor edad.

A raíz de esto es que se dividió esta variable de la siguiente manera;

- 1) 10 a 13 años
- 2) 14 a 18 años
- 3) 19 a 23 años
- 4) 24 a 28 años
- 5) 29 a 33 años
- 6) 34 a 38 años
- 7) 39 a 45 años
- 8) 46 a 50 años
- 9) 50 a 60 años
- 10) 61 años o más

Faltas

Se realizó un análisis descriptivo acerca de cómo se distribuían las faltas durante los años 2021, 2020 y 2019 y luego se definió como se podía agrupar esta variable.

Faltas justificadas – Y1

El estudiantado de educación primaria pública tiene entre 0 a 75 faltas justificadas durante todo el año 2021, y en promedio estas representan 0,72 faltas por estudiante.

Se analiza si en algún grupo particular de estudiantes, hay una diferencia en el promedio de faltas.

Se observa que en primer año los alumnos tendrían más faltas justificadas, pero no serían demasiadas en relación a los grados siguientes.

Grado	Promedio Y1_faltas_justificadas
1	0.76
2	0.75
3	0.70
4	0.72
5	0.73
6	0.71
Total	0,73

Tabla B 1. Faltas justificadas en Y1 según grado

Tampoco se observa una clara relación entre quienes se ubican en quintiles socioeconómicos más bajos que el resto. La diferencia entre estudiantes del quintil 1 y el 5 es pequeña.

Quintil	Promedio Y1_faltas_justificadas
1	0.60
2	0.78
3	0.81
4	0.91
5	0.57
Total	0.73

Tabla B 2. Faltas justificadas en Y1 según quintil

Donde sí se observa una diferencia importante en el promedio de faltas es entre estudiantes que asisten a centros en zonas rurales o urbanas. Niñas y niños de escuelas rurales tienen más faltas que quienes asisten a urbanas.

Zona	Promedio Y1_faltas_justificadas
Urbana	0.21
Rural	10.17
Total	0.73

Tabla B 3. Faltas justificadas en Y1 según zona

También observamos que estudiantes que asisten a jornadas de tiempo completo tienen más faltas que aquellos que no, pero quienes asisten a centros de tiempo extendido tienen en promedio menor faltas que el resto.

Turno	Promedio Y1_faltas_justificadas
Matutino	0.23
Vespertino	0.23
Extendido	0.00
Tiempo completo	10.25
Doble turno	0.34
Especial	0.26
Sin datos	1.44
Total	0.73

Tabla B 4. Faltas justificadas en Y1 según turno

Luego de esto, se analiza la cantidad de faltas de los alumnos y se divide la variable para 2021 en las siguientes categorías;

- 1) 0 faltas
- 2) 1 a 5 faltas
- 3) 6 a 10 faltas
- 4) 11 a 20 faltas
- 5) 21 a 30 faltas
- 6) 31 a 40 faltas
- 7) 41 a 50 faltas
- 8) 51 a 60 faltas
- 9) 61 a 70 faltas
- 10) 71 faltas o más

Cabe destacar que un estudiante concurre a su centro educativo 5 días a la semana, 20 días al mes, 180 días al año.

Faltas injustificadas – Y1

La matrícula de alumnos de primaria tiene entre 0 a 54 faltas justificadas durante todo el año 2021, y en promedio estas son 0,51 faltas por persona.

Se analiza si en algún grupo particular de estudiantes hay una diferencia en el promedio de faltas.

Llama la atención que en primer año se observan más faltas injustificadas que cuando avanza la vida escolar. Sin embargo, las faltas en primer año se distribuyen entre 0 y 46 pero entre segundo y sexto año, van de 0 a 54 faltas.

Grado	Promedio Y1_faltas_injustificadas
1	0.83
2	0.46
3	0.46
4	0.46
5	0.44
6	0.44
Total	0.52

Tabla B 5. Faltas injustificadas en Y1 según grado

A diferencia de lo visto con las faltas justificadas, en los quintiles socioeconómicos más bajos se observa un promedio de faltas más alto que entre quienes asisten a centros contextos socioeconómicos más favorables.

Quintil	Promedio Y1_faltas_injustificadas
1	0.70
2	0.60
3	0.51
4	0.44
5	0.36
Total	0.52

Tabla B 6. Faltas injustificadas en Y1 según quintil

En relación a si el centro es rural o urbano, no se identifican grandes diferencias en el promedio de faltas injustificadas entre estudiantes.

Zona	Promedio Y1_faltas_injustificadas
Urbana	0.52
Rural	0.47
Total	0.52

Tabla B 7. Faltas injustificadas en Y1 según zona

Por otro lado, al igual que lo visto en faltas justificadas, quienes asisten a tiempo parecerían tener menos faltas en promedio que el resto. Más allá de esto no se ven grandes diferencias en relación a la duración de los turnos.

Turno	Promedio Y1_faltas_injustificadas
Matutino	0.51
Vespertino	0.55
Extendido	0.20
Tiempo completo	0.48
Doble turno	0.50
Especial	0.42
Sin datos	0.53
Total	0.52

Tabla B 8. Faltas injustificadas en Y1 según turno

Se analiza la cantidad de faltas de los alumnos, y se divide la variable faltas injustificadas para 2021 en las siguientes categorías;

- 1) 0 faltas
- 2) 1 a 5 faltas
- 3) 6 a 10 faltas
- 4) 11 a 20 faltas
- 5) 21 a 30 faltas
- 6) 31 a 40 faltas
- 7) 41 a 50 faltas
- 8) 51 o más

Faltas totales – Y1

Los alumnos tienen entre 0 a 90 faltas totales en el 2021.

A partir de esto, se analiza la cantidad de faltas de los alumnos, y se divide la variable para 2021 en las siguientes categorías;

- 1) 0 faltas
- 2) 1 a 5 faltas
- 3) 6 a 10 faltas
- 4) 11 a 20 faltas
- 5) 21 a 30 faltas
- 6) 31 a 40 faltas
- 7) 41 a 50 faltas
- 8) 51 a 60 faltas
- 9) 61 a 70 faltas
- 10) 71 faltas o más

Faltas justificadas – Y2

Quiénes asisten a educación primaria durante el 2020 tienen entre 0 a 57 faltas y en promedio estas son 1,67 faltas por alumno. Cabe destacar que dada la situación de pandemia, hubo menos días con pasaje de lista en educación primaria.

En el año 2020 se incorpora al análisis, estudiantes en educación inicial, ya que se analiza la historia escolar de todos quienes cursaron educación primaria durante el 2021.

Se observa en promedio más faltas justificadas entre quienes asisten a nivel inicial 5 años y a primer año. Se observa una baja entre segundo y quinto año para volver a aumentar en sexto año de educación primaria.

Por otro lado, observamos que en nivel inicial y primer año, las faltas se distribuyen entre 0 a 57 y entre segundo en adelante varían de la misma manera.

Grado	Promedio Y2_faltas_justificadas
5 años – inicial	2.26
1	2.03
2	1.55
3	1.48
4	1.41
5	1.34
6	1.87
Total	1.67

Tabla B 9. Faltas justificadas en Y2 según grado

Se observa que ante niveles socioeconómicos más favorables, la matrícula mantiene una menor cantidad de falta justificadas. Sin embargo, entre estudiantes de quintil 1 las faltas van de 0 a 54, y entre los de quintil 5 se distribuyen de la misma manera.

Quintil	Promedio Y2_faltas_justificadas
1	1.94
2	1.86
3	1.73
4	1.61
5	1.30
Total	1.67

Tabla B 10. Faltas justificadas en Y2 según quintil

En relación a la ubicación del centro, no se identifican grandes diferencias entre quienes asisten a educación urbana o rural.

Zona	Promedio Y2_faltas_justificadas
Urbana	1.65
Rural	2.02
Total	1.67

Tabla B 11. Faltas justificadas en Y2 según zona

Por otro lado, se ve una mayor cantidad de faltas justificadas entre estudiantes de tiempo completo o de doble turno que en el resto. Nuevamente los alumnos de tiempo extendido mantienen un promedio menor de faltas justificadas

Turno	Promedio Y2_faltas_justificadas
Matutino	1.77
Vespertino	1.50
Extendido	0.59
Tiempo completo	2.05
Doble turno	1.83
Especial	1.33
Sin datos	1.50
Total	1.67

Tabla B 12. Faltas justificadas en Y2 según turno

A partir de esto, se analiza la cantidad de faltas de los alumnos, y se divide la variable para 2020 en las siguientes categorías;

- 9) 0 faltas
- 10) 1 a 5 faltas
- 11) 6 a 10 faltas
- 12) 11 a 20 faltas
- 13) 21 a 30 faltas
- 14) 31 a 40 faltas
- 15) 41 a 50 faltas
- 16) 51 o más

Faltas injustificadas – Y2

Durante todo el año 2020 se observan entre 0 a 111 faltas injustificadas, las cuales representan en promedio 4,67 faltas por estudiante.

En relación al grado se observa nuevamente que estudiantes de educación inicial, primer año y sexto año son quienes tienen más faltas.

Sin embargo, quienes asisten a educación inicial y primer año tienen entre 0 y 111 faltas pero quienes cursan sexto año tienen entre 0 a 56 faltas.

Grado	Promedio Y2_faltas_injustificadas
5 años	6.65
1	6.49
2	3.82
3	3.84
4	3.70
5	3.61
6	8.36
Total	4.67

Tabla B 13. Faltas injustificadas en Y2 según grado

En relación al nivel socioeconómico donde se ubica el centro educativo, se vuelve a observar un promedio de faltas más bajo ante contextos socioeconómicos menos vulnerables. Sin embargo, en los quintiles más bajos las faltas se distribuyen entre 0 y 111 y en los más altos entre 0 y 107.

Quintil	Promedio Y2_faltas_injustificadas
1	6.04
2	5.35
3	4.72
4	4.17
5	3.37
Total	4.67

Tabla B 14. Faltas injustificadas en Y2 según quintil

En relación a si el centro es rural o urbano, no se identifican grandes diferencias en el promedio de faltas injustificadas de los alumnos

Zona	Promedio Y2_faltas_injustificadas
Urbana	4.74
Rural	3.46
Total	4.67

Tabla B 15. Faltas injustificadas en Y2 según zona

Por otro lado, de forma distinta a lo que sucedía con las faltas justificadas, la parte del estudiantado que asiste tiempo completo, tiempo extendido y doble turno tienen menos faltas que quienes cursan en modalidad de turno simple (matutino o vespertino).

Turno	Promedio Y2_faltas_injustificadas
Matutino	4.98
Vespertino	4.78
Extendido	3.64
Tiempo completo	3.54
Doble turno	4.41
Especial	3.53
Sin datos	4.31
Total	4.67

Tabla B 16. Faltas injustificadas en Y2 según turno

Se divide entonces la variable faltas injustificadas para 2020 de la siguiente manera;

- 1) 0 faltas
- 2) 1 a 5 faltas
- 3) 6 a 10 faltas
- 4) 11 a 20 faltas
- 5) 21 a 30 faltas
- 6) 31 a 40 faltas
- 7) 41 a 50 faltas
- 8) 51 a 60 faltas
- 9) 61 a 70 faltas
- 10) 71 a 80 faltas
- 11) 81 faltas o más

Faltas totales – Y2

Los alumnos tienen entre 0 a 116 faltas totales en el 2020.

A partir de esto, se analiza la cantidad de faltas de los alumnos, y se divide la variable para 2020 en las siguientes categorías;

- 1) 0 faltas
- 2) 1 a 5 faltas
- 3) 6 a 10 faltas
- 4) 11 a 20 faltas
- 5) 21 a 30 faltas
- 6) 31 a 40 faltas
- 7) 41 a 50 faltas
- 8) 51 a 60 faltas
- 9) 61 a 70 faltas
- 10) 71 a 80 faltas
- 11) 81 faltas o más

Faltas justificadas – Y3

Durante el año 2019, el grupo de estudiantes analizado mantiene entre 0 a 180 faltas justificadas, las cuales representan en promedio 9,77 faltas por cada estudiante.

No se observa una clara disminución de las faltas a medida que avanza el año escolar como sucedía en años anteriores. Sin embargo en nivel 5 años, primer año y 5to año es donde se observan mayores faltas.

Grado	Promedio Y3_faltas_justificadas
4 años – inicial	5.82
5 años – inicial	12.79
1	11.57
2	9.83
3	9.33
4	8.95
5	10.25
6	7.73
Total	9.77

Tabla B 17. Faltas justificadas en Y3 según grado

Se observa que cuando el nivel socioeconómico donde se ubica el centro educativo es menos vulnerable, la cantidad de faltas es menor. Por otro lado, entre quienes asisten a centros de quintil 1 y 2 se observan entre 0 y 180 faltas, y entre quienes asisten a escuelas de quintil 4 y 5 las mismas se distribuyen entre 0 y 170.

Quintil	Promedio Y3_faltas_justificadas
1	10.52
2	10.40
3	10.36
4	9.50
5	8.40
Total	9.77

Tabla B 18. Faltas justificadas en Y3 según quintil

En relación a si el centro es rural o urbano, no se identifican grandes diferencias en el promedio de faltas justificadas para el año 2019.

Zona	Promedio Y3_faltas_justificadas
Urbana	9.75
Rural	10.30
Total	9.77

Tabla B 19. Faltas justificadas en Y3 según zona

Por otro lado, se observa también en este caso, que quienes cursan en centros de tiempo extendido tienen menos faltas que el resto del estudiantado.

Turno	Promedio Y3_faltas_justificadas
Matutino	10.28
Vespertino	9.21
Extendido	6.70
Tiempo completo	10.42
Doble turno	10.17
Especial	8.32
Sin datos	9.74
Total	9.77

Tabla B 20. Faltas justificadas en Y3 según turno

A partir de esto, se analiza la cantidad de faltas de los alumnos, y se divide la variable para 2021 en las siguientes categorías;

- 1) 0 faltas
- 2) 1 a 5 faltas
- 3) 6 a 10 faltas
- 4) 11 a 20 faltas
- 5) 21 a 30 faltas
- 6) 31 a 40 faltas
- 7) 41 a 50 faltas
- 8) 51 a 70 faltas
- 9) 71 a 90 faltas
- 10) 91 a 110 faltas
- 11) 111 a 130 faltas
- 12) 131 a 150 faltas
- 13) 151 a 170 faltas
- 14) 171 a 180 faltas

Faltas injustificadas – Y3

Durante el año 2019, la matrícula analizada tuvo entre 0 y 185 faltas injustificadas, las cuales significaron en promedio 16,91 faltas por estudiante.

En relación al grado, se observa que quienes asistieron a educación inicial, nivel 4 años fueron quienes tuvieron menos faltas, pero quienes cursaron nivel inicial en 5 años tuvieron una cantidad de faltas injustificadas mucho más alta.

Nuevamente vemos a los alumnos de sexto año como aquellos que más han faltado de manera injustificada.

Grado	Promedio Y3_faltas_injustificadas
4 años – inicial	9.62
5 años – inicial	24.75
1	23.65
2	14.72
3	14.05
4	13.71
5	30.13
6	55.45
Total	16.91

Tabla B 21. Faltas injustificadas en Y3 según grado

Se observa también en este caso que en los centros de mayor vulnerabilidad, el promedio de faltas injustificadas es más alto que en las escuelas de menor vulnerabilidad.

Quintil	Promedio Y3_faltas_injustificadas
1	22.32
2	19.00
3	16.87
4	15.44
5	11.98
Total	16.91

Tabla B 22. Faltas injustificadas en Y3 según quintil

En relación a la zona en que se ubica el centro educativo, quienes asisten a escuelas urbanas tienen en promedio seis faltas más que quienes cursan en escuelas pertenecientes a zonas rurales.

Zona	Promedio Y3_faltas_injustificadas
Urbana	17.19
Rural	11.69
Total	16.91

Tabla B 23. Faltas injustificadas en Y3 según zona

A diferencia de lo observado en años anteriores, los estudiantes que asisten a escuelas de tiempo completo tienen un promedio menor de faltas. Por otro lado, quienes cursan en modalidades simples (matutino o vespertino) tienen una mayor cantidad de faltas.

Turno	Promedio Y3_faltas_injustificadas
Matutino	18.13
Vespertino	17.66
Extendido	15.03
Tiempo completo	11.81
Doble turno	14.98
Especial	12.91
Sin datos	16.51
Total	16.91

Tabla B 24. Faltas injustificadas en Y3 según turno

A partir de esto, se analiza la cantidad de faltas de los alumnos, y se divide la variable para 2019 en las siguientes categorías;

- 1) 0 faltas
- 2) 1 a 5 faltas
- 3) 6 a 10 faltas
- 4) 11 a 20 faltas
- 5) 21 a 30 faltas
- 6) 31 a 40 faltas
- 7) 41 a 50 faltas
- 8) 51 a 70 faltas
- 9) 71 a 90 faltas
- 10) 91 a 110 faltas
- 11) 111 a 130 faltas
- 12) 131 a 150 faltas
- 13) 151 a 170 faltas
- 14) 171 o más

Faltas totales – Y3

La matrícula estudiantil tiene en 2019 entre 0 a 185 faltas totales. Se divide entonces a las faltas totales de la siguiente manera;

- 1) 0 faltas
- 2) 1 a 5 faltas
- 3) 6 a 10 faltas
- 4) 11 a 20 faltas
- 5) 21 a 30 faltas
- 6) 31 a 40 faltas
- 7) 41 a 50 faltas
- 8) 51 a 70 faltas
- 9) 71 a 90 faltas
- 10) 91 a 110 faltas
- 11) 111 a 130 faltas
- 12) 131 a 150 faltas
- 13) 151 a 170 faltas
- 14) 171 a 180 faltas
- 15) 180 o más