



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL

IMPUTACIÓN DE LOS RESULTADOS DE UNA ENCUESTA A LOS
ESTABLECIMIENTOS EDUCACIONALES DE CHILE PARA PRIORIZAR
INTERVENCIONES EN CONVIVENCIA ESCOLAR

TESIS PARA OPTAR AL GRADO DE MAGÍSTER EN GESTIÓN Y POLÍTICAS
PÚBLICAS

MARIO ANTONIO FUENTES ELGUETA

PROFESOR GUÍA:
PATRICIO RODRÍGUEZ VALDÉS
PROFESORA CO-GUÍA:
VERÓNICA LÓPEZ LEIVA

MIEMBRO DE LA COMISIÓN:
SERGIO CELIS GUZMÁN

SANTIAGO DE CHILE
2021

RESUMEN DE LA TESIS PARA OPTAR AL GRADO DE:
Magíster en Gestión y Políticas Públicas.
POR: Mario Fuentes Elgueta.
FECHA: Diciembre 2021.
PROFESOR GUÍA: Patricio Rodríguez Valdés.
PROFESORA CO-GUÍA: Verónica López Leiva.

IMPUTACIÓN DE LOS RESULTADOS DE UNA ENCUESTA A LOS ESTABLECIMIENTOS EDUCACIONALES DE CHILE PARA PRIORIZAR INTERVENCIONES EN CONVIVENCIA ESCOLAR

La convivencia escolar es un fenómeno multidimensional con grandes repercusiones en los estudiantes, las escuelas y la sociedad. La literatura experta destaca su importancia para obtener mejores niveles de rendimiento académico, menores índices de deserción escolar y el bienestar emocional de las comunidades educativas. También ha destacado su relevancia para la convivencia social, a través de la generación de una ciudadanía activa, crítica y participativa.

El Programa PACES-PUCV desarrolló un Sistema de Monitoreo de la Convivencia Escolar con el propósito de medir, analizar y gestionar la convivencia en las escuelas. Este sistema consta de una encuesta, emisión de reportes y asistencia técnica para los equipos de mejoramiento de la convivencia, recoge evidencia de experiencias internacionales, las adapta al contexto nacional y se sustenta en el modelo social-ecológico, que diagnostica holísticamente la convivencia escolar. Sin embargo, el Sistema de Monitoreo actualmente posee una cobertura limitada.

Es por ello que el objetivo principal del presente estudio es estimar el resultado de los establecimientos educacionales del país en la Encuesta PACES-PUCV, lo que permitirá seleccionar establecimientos educativos a priorizar en programas de mejoramiento de la convivencia escolar. Este objetivo se alcanzó mediante algoritmos de aprendizaje automático, que permitieron clasificar las escuelas del país según su tipo de convivencia.

Gracias a los algoritmos de *machine learning* y una muestra de escuelas PACES-PUCV (n=1346) se consiguió una clasificación binaria nacional de escuelas según su “débil” o “buena” convivencia escolar. El mejor modelo de clasificación obtuvo un *accuracy* del 78,51%, logrando una *precisión* del 66,1% y un *recall* del 81,25% para la clase “débil” convivencia, lo que muestra su capacidad para identificar y clasificar correctamente ambas categorías de escuela, y disminuir a la vez los “falsos negativos”.

Lo anterior permite concluir que es posible imputar el resultado la Encuesta PACES-PUCV. Dado que esta clasificación representa el riesgo de la escuela de presentar problemas de convivencia, los resultados permiten remediar una brecha información y focalizar los esfuerzos de la política pública en escuelas que requieren fortalecer la gestión de la convivencia. Por último, esta investigación abre nuevas áreas de indagación, como por ejemplo, realizar una clasificación multiclase que permita comprender de forma más detallada la convivencia a nivel nacional.

Agradecimientos¹

Agradezco a mi Profesor Guía, Patricio Rodríguez Valdés, por sus valiosos consejos y orientaciones a lo largo de todo el proceso de investigación.

Agradezco también a mi Profesora Co-Guía, Verónica López Leiva y al equipo del Programa PACES-PUCV, por su tiempo, orientaciones en materia de convivencia escolar y el imprescindible acceso que me entregaron a los datos para realizar este trabajo.

A mi familia, amigos y pareja, por su compañía incondicional.

¹ Esta tesis se realizó en el marco del Proyecto Fondecyt Regular N°1191267 “La inclusión excluyente: Prácticas de segregación y exclusión en escuelas chilenas”, en el cual la Profesora Verónica López Leiva es Investigadora Principal.

Tabla de Contenido

1	Planteamiento del problema de investigación	1
1.1	Introducción	1
1.2	Estructura de la tesis	3
1.3	Problema de investigación	3
1.4	El análisis de datos para apoyar la toma de decisiones	4
1.5	Pregunta y objetivos de investigación	4
1.5.1	Objetivo general	6
1.5.2	Objetivos específicos	6
2	Revisión de literatura	7
2.1	¿Qué es la convivencia escolar?	7
2.2	¿Qué factores influyen en la convivencia escolar?	10
2.2.1	Factores individuales	11
2.2.2	Factores familiares	11
2.2.3	Factores escolares	12
2.2.4	Factores contextuales	13
2.2.5	Factores de política pública	13
2.3	Medidas de la convivencia escolar en Chile	15
2.3.1	Indicador Clima de Convivencia Escolar	15
2.3.2	¿Un indicador desde el modelo de la Encuesta PACES-PUCV?	17
2.4	Predicción y aprendizaje automático	19
3	Marco Metodológico	23
3.1	Exploración de datos	26
3.2	Preparación de datos	28
3.2.1	Cálculo de resultados PACES a nivel de establecimiento	28
3.2.2	Construcción de base de datos	28
3.2.3	Análisis exploratorio de datos	29
3.3	Planificación del modelo	32

3.4	Elaboración del modelo	32
3.5	Comunicar resultados	32
3.6	Uso	32
4	Resultados	33
4.1	Modelos de Regresión	33
4.2	Modelos de Clasificación	34
4.2.1	Resumen y evaluación de modelos de clasificación	35
4.3	Importancia de las variables para el modelo de clasificación	38
4.4	Clasificación nacional usando el modelo de Regresión Logística	40
5	Limitaciones, conclusiones y trabajo futuro	44
5.1	Limitaciones	44
5.2	Conclusiones	46
5.3	Trabajo futuro	49
6.	Bibliografía	51

Índice de Tablas

Tabla N°1.	Resumen comparativo de conceptos de convivencia escolar	10
Tabla N°2.	Resumen factores que influyen la convivencia escolar	14
Tabla N°3.	Encuesta IDPS Clima de Convivencia Escolar 6b 2016	15
Tabla N°4.	Cuadro comparativo de mediciones de convivencia escolar	18
Tabla N°5.	Fuentes de datos utilizadas	26
Tabla N°6.	VARIABLES SIMCE 2018-2016 y sus equivalencias	27
Tabla N°7.	VARIABLES utilizadas Cuestionarios de Calidad y Contexto SIMCE 2016	30
Tabla N°8.	VARIABLES utilizadas Cuestionarios de Calidad y Contexto SIMCE 2018	31
Tabla N°9.	Resultado modelos de regresión “clima de aula”, “clima escolar”	34
Tabla N°10.	Resultado modelos de regresión “clima de aula”, “clima escolar” (R ² ajustados)	34
Tabla N°11.	Resultados por clase con validación cruzada (\bar{x})	35

Tabla N°12. Resultados por clase con validación cruzada (\bar{x} menos $\sigma=0,2$).....	36
Tabla N°13. Resultados por clase con validación cruzada (\bar{x} menos $\sigma=0,3$).....	38
Tabla N°14. Resultados por clase con validación cruzada (\bar{x} menos $\sigma=0,4$).....	38
Tabla N°15. Importancia de las variables para el modelo Regresión Logística.....	39
Tabla N°16. Porcentaje de escuelas según tipo de convivencia y región.....	40

Índice de Figuras

Figura N°1. Proceso de elaboración del ICCE.	17
Figura N°2. Funcionamiento del aprendizaje supervisado.	20
Figura N°3. Funcionamiento del aprendizaje reforzado.....	21
Figura N°4. Tipos de Aprendizaje Automático.....	22
Figura N°5. Ciclo de vida del análisis de datos.....	23
Figura N°6. Matriz de confusión.....	37

1 Planteamiento del problema de investigación

1.1 Introducción

La convivencia escolar tiene una importancia decisiva para el desarrollo de los procesos de enseñanza-aprendizaje y el bienestar emocional de las comunidades educativas (Agencia de Calidad de la Educación, 2017; Mineduc, 2019a; Mineduc 2019b). Existe abundante evidencia de que un clima escolar positivo promueve el aprendizaje y está asociado con mayores logros académicos de los estudiantes² (UNESCO, 2008; Konold, Cornell, Jia & Malone, 2018). La literatura experta subraya que las escuelas con un clima escolar positivo presentan mejores niveles de rendimiento y menores índices de deserción escolar (Paredes & Frazer, 1992; Cheng, 1994). Además, investigadores han destacado que mejorar la convivencia escolar es un fin valioso en sí mismo, que aporta a una sociedad más justa y participativa (Ascorra y López, 2019) y es un elemento decisivo en la configuración de la convivencia social y la generación de una ciudadanía activa y participativa (Tuvilla, 2004; Perales, Arias & Bazdresch, 2014). Es por esto que contar con mediciones robustas, claras y completas de la convivencia escolar es primordial para que los tomadores de decisiones y las comunidades educativas puedan trabajar en su mejora.

Según la Historia de la Ley N° 20.529, de 2011, con esta normativa se buscaba generar mecanismos de aseguramiento que protegieran el derecho de los ciudadanos a recibir una educación de calidad (BCN, 2021). Por medio de su promulgación se modificó la arquitectura institucional del sistema educativo nacional, creándose dos nuevos organismos públicos: la Agencia de Calidad de la Educación (ACE) y la Superintendencia de Educación (Supereduc). La ACE está a cargo de evaluar y orientar al sistema educativo para que propenda al mejoramiento de la calidad y equidad de las oportunidades educativas, mientras que la Supereduc es la encargada de la fiscalización del cumplimiento de la normativa educacional y de la legalidad del uso de recursos que reciban los sostenedores receptores de aportes estatales. Al mismo tiempo, la norma citada creó el Sistema de Aseguramiento de la Calidad de la Educación (SAC) que, con el Ministerio de Educación (Mineduc) como órgano rector, tiene como propósito asegurar el acceso a una educación de calidad con equidad para todos los estudiantes del país.

La ACE es un servicio público descentralizado, con consejo exclusivo, dotado de personalidad jurídica y patrimonio propio, que se relaciona con el presidente de la República por medio del Mineduc. Entre sus funciones se encuentran evaluar los logros de aprendizaje de las escuelas, informar sobre los resultados de sus evaluaciones y orientar el desempeño de establecimientos, directores y sostenedores. Esta institución entrega retroalimentación periódica a las escuelas del

² Se utiliza “clima escolar” como sinónimo de “convivencia escolar” aunque son conceptos relacionados pero diferentes. La convivencia escolar se refiere a prácticas concretas, cotidianas y situadas localmente que se producen de forma sostenida en el tiempo, mientras que el clima escolar se refiere a la percepción de los estudiantes o la comunidad escolar sobre la calidad del ambiente escolar (López, 2020).

país sobre sus resultados educativos, lo que incluye sus resultados en mediciones de rendimiento académico (SIMCE) así como resultados en dimensiones no académicas de la calidad educativa, como las contempladas en los Indicadores de Desarrollo Personal y Social (IDPS). De esta forma, los IDPS amplían el concepto de calidad de la educación, al complementar los cálculos tradicionales de rendimiento académico con la medición de índices sobre autoestima académica y motivación escolar, participación y formación ciudadana, hábitos de vida saludable, clima de convivencia escolar, entre algunos otros (ACE, 2018). Los resultados de estas evaluaciones se informan periódicamente a la ciudadanía y a los establecimientos educacionales del país.

El indicador de la ACE diseñado para medir el clima de convivencia en las escuelas se denomina “Indicador de Clima de Convivencia Escolar” (ICCE). El ICCE intenta revelar cómo actúan y perciben la convivencia escolar estudiantes, docentes, padres y apoderados, definiéndola como producto de un ambiente de respeto, organizado y seguro (ACE, 2017). Un ambiente de respeto considera las percepciones y actitudes en relación con el trato respetuoso entre los miembros de la comunidad educativa, la valoración de la diversidad y la ausencia de discriminación. Un ambiente organizado considera las percepciones y actitudes sobre la existencia de normas claras, conocidas, exigidas y respetadas por todos y el predominio de mecanismos constructivos de resolución de conflictos. Finalmente, un ambiente seguro alude a las percepciones y actitudes con relación al grado de seguridad y de violencia física y psicológica al interior del establecimiento, así como la existencia de mecanismos de prevención ante ésta (ACE, 2016).

Por su parte, el Programa de Apoyo a la Convivencia Escolar de la Pontificia Universidad Católica de Valparaíso (Programa PACES-PUCV) creó un Sistema de Monitoreo de la Convivencia Escolar que fue transferido a las escuelas beneficiarias del Programa Habilidades para la Vida II (HPV II) de Junta Nacional de Auxilio Escolar y Becas (JUNAEB), a fin de medir, analizar y gestionar su convivencia. Este sistema está compuesto por una encuesta, emisión de reportes descargables por cada escuela y asistencia técnica para el equipo de gestión escolar en mejoramiento de la convivencia (Ascorra y López, 2019). Su construcción recogió evidencia de experiencias internacionales para el diseño de sistemas de monitoreo de la convivencia escolar, y los adaptó al contexto nacional, con la finalidad de proveer información confiable e interpretable para las escuelas, con una visión de mayor complejidad en relación a los aspectos a considerar (López, Bilbao, Ascorra, Carrasco, Benbenishty, Astor & Álvarez, 2019).

En particular, y teniendo como base teórica un modelo social-ecológico³, la Encuesta PACES considera una serie de variables agrupadas en subdimensiones y escalas que permiten diagnosticar holísticamente la situación de la convivencia en una escuela. Por ejemplo, la escala de “clima escolar”, considera una agrupación de variables asociadas a la satisfacción del

³ Donde se entiende que el fenómeno de la convivencia escolar está inserto dentro de las dinámicas relacionales de la escuela, que guardan relación con el clima de los espacios de aprendizaje y convivencia, así como con aspectos del entorno y de políticas del establecimiento educacional (Astor & Benbenishty, 2019, citado por Ascorra y López, 2019).

estudiante con la escuela, el apoyo que percibe y la forma en que observa la participación estudiantil en el establecimiento. La escala “clima de aula” considera su percepción del ambiente físico de la escuela, la organización y participación del curso, las relaciones entre pares, junto a la orientación al aprendizaje de los alumnos. Como un último ejemplo, la escala de “victimización entre pares” recoge información sobre victimización verbal, por medios electrónicos, física, por exclusión social y por acoso sexual. De esta manera, el instrumento creado por PACES provee una mirada holística de la convivencia en la escuela, lo cual posibilita intervenciones para su mejoramiento por parte de la comunidad del establecimiento.

Este instrumento actualmente posee una cobertura limitada, que está circunscrita a las escuelas participantes del programa Habilidades para la Vida (HPV II) JUNAEB. Es por ello que el objetivo principal del presente estudio es estimar el resultado general de los establecimientos educacionales del país en la Encuesta PACES. Esta estimación permitiría seleccionar establecimientos educativos a priorizar, a fin de escalar el programa HPV II, en aquellas escuelas que muestren un peor pronóstico de resultados en la medición.

1.2 Estructura de la tesis

Esta tesis se estructura como sigue. En el primer capítulo se plantea el problema de investigación, relevando la importancia individual, para las escuelas y para la sociedad de la medición y buena gestión de la convivencia escolar. El segundo capítulo, revisa y presenta literatura sobre convivencia escolar, cómo se define y qué factores la influyen, describiendo además los instrumentos que la miden en el país. El mismo capítulo, además, presenta los tipos de aprendizaje automático y sus características generales. El tercer capítulo expone en general la metodología del estudio, describe y muestra su aplicación paso a paso. En el cuarto capítulo se crea una clasificación de escuelas en aquellas con “débil” o “buena” convivencia escolar, mostrándose los resultados más importantes en materia de regresión y clasificación una vez entrenados los modelos, junto a sus criterios de evaluación. En el mismo apartado, y en base al modelo con mejores resultados, se realiza una imputación a 6.820 escuelas del país. Finalmente, el quinto capítulo expone las principales limitaciones del estudio, las conclusiones que se desprenden de este trabajo y sugiere investigaciones posteriores, que considera complementarias a la desarrollada en este estudio teórico-práctico.

1.3 Problema de investigación

La Encuesta PACES, creada por el Programa PACES-PUCV es una forma de medición de la convivencia escolar. Ésta considera la perspectiva de estudiantes, docentes y apoderados por separado (Ramírez, Ascorra, Carrasco-Aguilar & López, 2019), recoge investigación y evidencia internacional en una serie de dimensiones que abarcan clima escolar, victimización, satisfacción con la escuela, peligrosidad en la escuela y barrio, bienestar subjetivo, entre otras detalladas más adelante, lo que implica abordar la convivencia desde una perspectiva teórica multidimensional y un modelo ecológico, que considera capas individuales, organizacionales y comunitarias del problema. Esta medición es parte del Sistema de Monitoreo de la Convivencia Escolar de

PACES, que es utilizado por la JUNAEB en su Programa HPV II, y que el año 2018 fue aplicado en más de 1.000 escuelas a lo largo del país, abarcando aproximadamente 108.000 estudiantes.

La Encuesta PACES cuenta en este momento con una limitada cobertura y busca ampliarse dentro del país, como parte del programa HPV II. Una forma de identificar en qué establecimientos educacionales expandir el programa HPV II es estimar sus resultados en la Encuesta PACES de convivencia escolar, que será el instrumento con el que serán diagnosticados, para luego priorizar aquellas escuelas que tengan un peor pronóstico.

De ahí surge la pregunta ¿es posible imputar los resultados de la Encuesta PACES a establecimientos educacionales que no estén en el programa usando datos administrativos del sector público? Realizar esta imputación contribuiría a cerrar una brecha de información para la toma de decisiones de política educativa mostrando niveles de riesgo de las escuelas en las dimensiones o subdimensiones de su convivencia escolar, y clasificando los establecimientos según su “buena” o “débil” convivencia para determinar en qué otros establecimientos implementar el programa HPV II.

1.4 El análisis de datos para apoyar la toma de decisiones

El potencial de los datos estructurados y no estructurados para apoyar la toma de decisiones en el sector público y privado crece diariamente por su proliferación, la enorme capacidad de procesamiento disponible y la disminución de costes de almacenamiento. Una gran cantidad de autores han destacado la importancia del uso de evidencia y datos para el diseño e implementación de las políticas públicas en áreas tan diversas como ciudades inteligentes, seguridad, declaración de renta, salud, delincuencia, educación y muchos otros (Tomar et al., 2016; Siegel, 2013). Otros investigadores han mostrado su aplicación a áreas de política social, como la política educativa (Rodríguez, Suchan, Truffello & Matas, 2016; Rodríguez, Palomino & Mondaca, 2017). Sin duda las aplicaciones en el sector privado han aumentado en áreas como el *retail*, la industria del entretenimiento, la banca, la bolsa, el tráfico vehicular, rendimiento laboral, entre muchos otros temas (Siegel, 2013; Davenport, 2016).

Dentro de las técnicas de análisis de datos, el aprendizaje automático ha sido definido como una subespecialidad de la ciencia de la computación que se ocupa del diseño y desarrollo de algoritmos que permiten inferir comportamientos en base a datos empíricos (Rodríguez, Palomino & Mondaca, 2017) o la aplicación y ciencia de los algoritmos que da sentido a los datos, para convertir estos datos en conocimiento (Raschka & Mirjalili, 2019), a través de los métodos y algoritmos utilizados.

1.5 Pregunta y objetivos de investigación

Considerando que el uso de la intervención de HPV II podría traducirse en potenciales efectos positivos para las escuelas, cabe preguntarse si es posible imputar el resultado de la medición de la Encuesta PACES para identificar qué escuelas incorporar con mayor urgencia al programa.

Este trabajo busca resolver el problema por medio de la aplicación del aprendizaje automático o “*machine learning*”. A través de algoritmos de regresión o clasificación se conseguirá una estimación de los valores que obtendrían las escuelas en la Encuesta PACES si se les aplicara dicho instrumento, o bien una clasificación de ellos según una tipología de convivencia escolar⁴.

La presente investigación aplicada pretende contribuir al estudio de la convivencia escolar desde el punto de vista de la analítica predictiva. La relevancia de este estudio está en explorar la posibilidad de hacer predicciones sobre la convivencia escolar de las escuelas, desde el modelo del Programa PACES-PUCV, y en base a datos administrativos del sistema educativo chileno. A su vez, se considera relevante contribuir al asentamiento nuevas metodologías de análisis de datos para la toma de decisiones basada en evidencia. Esta exploración resulta urgente en un contexto donde, según la Superintendencia de Educación (2020), aproximadamente el 70% de las denuncias que recibe anualmente tienen relación con temas de convivencia escolar.

Ante la limitada cobertura actual de la Encuesta PACES, esta investigación entregará un insumo para la toma de decisiones de política pública. Aportar este insumo se justifica por las propiedades del instrumento de recogida de información del Programa PACES-PUCV, que aborda la convivencia escolar holísticamente, desde una perspectiva teórica multidimensional y ecológica, lo que permitiría cerrar una brecha de información para la toma de decisiones. Concretamente, la investigación planteada contribuirá a resolver este problema aplicando modelos de aprendizaje automático que permitan saber qué establecimientos podrían ser parte de un grupo de riesgo en cuanto a problemas de convivencia escolar. Este instrumento podría mejorar la focalización de políticas y programas de convivencia escolar generados por el Ministerio de Educación, por medio de la identificación de tipologías de convivencia en los establecimientos (“buena” o “débil” convivencia, por ejemplo), que permitan adaptar las intervenciones del nivel central a las realidades territoriales.

Para el nivel meso o de Servicios Locales de Educación, este instrumento permitiría profundizar el conocimiento existente sobre las realidades locales, identificar aquellos sectores donde los problemas de violencia escolar y seguridad en la escuela podrían ser preponderantes, a fin de coordinar acciones conjuntas con otras instituciones públicas. Al tratarse de problemas complejos, especialmente cuando se los observa bajo una mirada “ecológica”, se requiere una intervención capaz de abordar el problema con la misma complejidad. Por último, a nivel local o micro, esta información serviría para generar intervenciones específicas, en escuelas y aulas. Claramente la mayor efectividad de las intervenciones ocurriría cuando los distintos niveles institucionales comparten su información y generan intervenciones coordinadas y con un propósito común.

⁴ En el Capítulo N°4 se propone una clasificación entre escuelas con “buena” y “débil” convivencia escolar.

1.5.1 Objetivo general

Predecir el resultado de establecimientos educacionales en mediciones PACES que no han respondido la encuesta, a través de un modelo de aprendizaje automático a partir de datos administrativos del sector público y los resultados de los establecimientos que si la contestaron.

1.5.2 Objetivos específicos

- 1) Revisar en la literatura especializada los distintos conceptos de convivencia escolar que están presentes.
- 2) A partir de la literatura especializada en el tema analizar las variables administrativas que pueden identificar problemas de convivencia escolar, en base al modelo de PACES y cuyos datos sea posible obtener a partir de fuentes administrativas.
- 3) Comparar la capacidad de distintos modelos de aprendizaje supervisado de predecir el resultado de la encuesta (o indicador) minimizando el posible error de imputación.
- 4) Imputar el resultado del indicador de convivencia escolar PACES a aquellos establecimientos que no hayan contestado la Encuesta de Convivencia Escolar.

2 Revisión de literatura

2.1 ¿Qué es la convivencia escolar?

El material para redactar este apartado se obtuvo de la revisión de literatura especializada sobre “convivencia escolar” en los últimos veinte años. Los criterios de selección de los artículos o libros fueron que recopilaran definiciones del concepto o formularan una definición propia y que fueran desarrollados en Latinoamérica, lo que limitó la búsqueda principalmente a textos escritos en español. El resultado fue la identificación y exposición de tres conceptos diferentes de convivencia en las escuelas, a partir de más de diez artículos o libros seleccionados.

Como elementos descriptivos transversalmente presentes en la literatura revisada, destaca la comprensión de la convivencia como una construcción colectiva (Mineduc 2019b; Banz, 2008; Tuvilla, 2004) cuya responsabilidad recae en todos los miembros de las comunidades escolares (Mineduc 2019b; Banz, 2008; Tuvilla, 2004) y con una importancia decisiva para el desarrollo de procesos de enseñanza-aprendizaje y el bienestar emocional de los actores de las comunidades educativas (Agencia de Calidad de la Educación, 2017; Mineduc, 2019a; Mineduc 2019b).

Y aunque se aprecien rasgos transversales en la conceptualización de la convivencia escolar entre los especialistas, también es habitual encontrar referencias a su carácter polisémico (Fierro y Carbajal, 2019; Morales y López, 2019), multidimensional y elusivo (Perales, Arias y Bazdresch, 2014), utilizado de forma tan diversa que lleva a la confusión (Fierro y Carbajal, 2019) y a la proliferación de conceptos que emergen de diferentes paradigmas. En esa línea, algunos autores han identificado aproximaciones de carácter positivista al concepto (Ascorra et al., 2017) y otros han reconocido conceptos desarrollados desde la teoría crítica (Magendzo et al., 2013).

Para la Política Nacional de Convivencia Escolar (PNCE) *“la convivencia escolar es el conjunto de las interacciones y relaciones que se producen entre todos los actores de la comunidad”* (Mineduc, 2019:9). Una caracterización similar presenta López (2020), quien sostiene que en la literatura iberoamericana la convivencia escolar se refiere a prácticas concretas, cotidianas y situadas localmente que se producen de forma sostenida en el tiempo, entre personas y grupos con diferencias socioeconómicas, étnicas, de género, entre otras⁵. En la tradición descrita por esta autora, el concepto de convivencia escolar *“está más vinculado a los valores culturales de fortalecimiento de la democracia y desarrollar capacidades para la ciudadanía”* (López, 2020:58). Esta mirada encaminada al “saber vivir con otros” se encuentra también explícita en la PNCE, donde se señala que la convivencia es clave para el desarrollo social de los miembros de las comunidades escolares *“porque permite construir una ciudadanía comprometida con el*

⁵ La convivencia escolar se distingue del “clima escolar” pues éste se refiere a la percepción de los estudiantes o la comunidad escolar sobre la calidad del ambiente escolar (López, 2020). Una definición similar ofrece la PNCE, que caracteriza el clima escolar como las “percepciones que tienen los sujetos de las relaciones y de las condiciones del ambiente en el que estas se producen” (Mineduc, 2020:12).

cuidado de los demás, de los bienes públicos y de la democracia” (Mineduc, 2019:11). En esta perspectiva, las fricciones en las relaciones de las personas y grupos de las comunidades escolares pueden tener un valor formativo, alcanzado a través de la participación, co-creación de diagnósticos y co-construcción de las alternativas de solución.

En aquellos puntos de vista convive una descripción de la convivencia y una noción específica sobre la finalidad de ésta en la escuela y, en una perspectiva mucho más amplia, de ésta para la sociedad. La gestión de la convivencia se caracterizará, entonces, por estrategias escolares de tipo formativas y normas construidas de forma participativa y horizontal.

Un segundo punto de vista, sostenido por Perales, Arias y Bazdresch (2014) argumenta que la convivencia es reflejo de la cultura que se desarrolla en la vida cotidiana, que implica el modo de ser de cada persona en interrelación con los otros, donde cada sujeto pone en práctica sus valores, creencias, percepciones, reflexiones, el conocimiento adquirido, entre otros. Este es un enfoque centrado en el individuo, su autonomía, autoeficacia, capacidad de reflexión, sentido de propósito, autorregulación y mentalidad de crecimiento. Por medio del autoconocimiento, la confianza, el entendimiento y la regulación de sus emociones, es posible para el individuo generar relaciones sociales armónicas, positivas, responsables y con sentido de comunidad. Es así como esta perspectiva releva la importancia de las habilidades socioemocionales y el desarrollo de los recursos cognitivos y emocionales individuales para enfrentar situaciones diarias y ser capaces de encausar los conflictos en los establecimientos educativos.

Al igual que en la perspectiva de la PNCE, desde esta óptica *“la convivencia es una meta esencial de la educación que no se puede reducir al ámbito exclusivo de aula, sino que está presente en todos los espacios, mediatos e inmediatos, que conforman la escuela”* (Perales, Arias & Bazdresch, 2014:17). Además, destaca que la experiencia en la escuela puede contribuir en la posibilidad de experimentar lo que puede significar una ciudadanía activa y proactiva, participativa y deliberante (Perales, Arias & Bazdresch, 2014). De esta forma, si bien existen similitudes entre ambas perspectivas, también se constatan elementos que predominan sobre otros, ya que, para Perales, Arias & Bazdresch (2014) se busca gestionar las emociones para autorregular las conductas y propiciar que las emociones se manifiesten constructivamente. En cambio, desde la mirada de la PNCE el foco se encuentra más orientado a construir una visión compartida de la convivencia que queremos como comunidad.

Desde un tercer punto de vista, la convivencia escolar es un instrumento para evitar la violencia y está centrada especialmente en la construcción de espacios seguros, disciplinados y ordenados, con tolerancia cero a la violencia. De esta forma, para Carrasco, López y Estay (2012), la Ley N° 20.536, de Violencia Escolar, comprendió la convivencia escolar como un factor situacional inhibidor de violencia, en que la responsabilidad de la buena convivencia escolar recae en los adultos trabajadores de la escuela, invisibilizando la co-construcción y, por tanto, la co-responsabilidad de alumnos y apoderados (Carrasco, López & Estay, 2012). Por medio de un análisis crítico documental, las autoras argumentan que esta ley trae consigo el ejercicio de un poder disciplinario y la prescripción de una vigilancia jerárquica por parte de los adultos que trabajan en el establecimiento educacional (Carrasco, López & Estay, 2012).

En una posición crítica similar Magendzo, Toledo y Gutiérrez (2013) argumentan que en la Ley N° 20.536 conviven dos paradigmas antagónicos: a) control y sanción; b) convivencia escolar democrática. El primero se funda en la postura de seguridad nacional y corresponde a una concepción conductista de la educación donde las interacciones sociales se regulan a través del reforzamiento y el castigo, mientras que el segundo se funda en la concepción de seguridad humana y corresponde a una concepción crítica de la educación que apunta hacia la reconstrucción social. Al paradigma de control y sanción subyace el supuesto de que si se mitigan o erradican las conductas violentas aumentará la buena convivencia, lo cual, a juicio de los autores, no cuenta con un soporte empírico que lo avale (Magendzo, Toledo y Gutiérrez, 2013).

Esta mirada apunta a eliminar la violencia y sancionar a quienes la ejercen, con políticas de tolerancia cero a la violencia y la indisciplina. No obstante, *“naturalizar un hecho conflictivo en algunos estudiantes rompe con la perspectiva relacional de la convivencia escolar, saca el problema de la escuela y puede generar procesos de exclusión, concentrándose en que esos estudiantes cambien sus conductas y dejando el resto de las interacciones sin cuestionar”* (Perales, Arias & Bazdresch, 2014:59) y, como señalan López, Carrasco, Morales, Ayala, Karmy, y López (2011) esta “individualización de violencia escolar” comprende la violencia escolar como un problema individual, percibiendo al alumno como problemático debido a sus rasgos psicológicos y/o debido a familias negligentes. Ello a la larga se asociaría con prácticas de “externalización de responsabilidades” de la escuela frente a la violencia que emerge en sus aulas.

De esta manera, se observa que no existe consenso sobre qué es la convivencia escolar, sus finalidades y modo de gestionarla en las escuelas. Sin embargo, existen rasgos transversales en las definiciones de convivencia escolar, como la preocupación por el conflicto en las escuelas y el *bullying*. Existe también cierta complementariedad en algunos de esos planteamientos teóricos, por ejemplo, en la convivencia como “preparación para la democracia” y como “educación socio emocional o socioafectiva”, mientras que la mirada “de seguridad” está más centrada en la búsqueda de seguridad, orden y disciplina dentro de los espacios escolares.

La Tabla N° 1 resume los tres conceptos de convivencia escolar, que también expresan una idea sobre la gestión de la convivencia escolar: a) como preparación para la democracia desde un enfoque de derechos; b) como educación socioemocional y c) como evitación de la violencia o de seguridad. Estas formas de comprender la convivencia pueden distinguirse a su vez por su foco en: a) individuos; b) grupos de individuos; c) relaciones entre esos individuos y grupos:

Tabla N°1. Resumen comparativo de conceptos de convivencia escolar.

Ítem	Como preparación para la democracia.	Como educación socioemocional	Como evitación de la violencia o de seguridad
¿Qué es?	La convivencia escolar es el conjunto de las interacciones y relaciones que se producen entre todos los actores de la comunidad, abarcando no solo aquellas entre individuos, sino también las que se producen entre los grupos, equipos, cursos y organizaciones internas que forman parte de la institución. Incluye también la relación de la comunidad escolar con las organizaciones del entorno en el que se encuentra inserta (Mineduc, 2019b)	La convivencia es reflejo de la cultura que se desarrolla en la vida cotidiana, implica el modo de ser de cada persona en interrelación con los otros, donde cada sujeto pone en práctica sus valores, creencias, percepciones, reflexiones, el conocimiento adquirido, etc. (Perales, Arias & Bazdresch, 2014)	Comprende la convivencia como “la coexistencia armónica de los miembros de la comunidad educativa, que supone una interrelación positiva entre ellos”
Autores	Ascorra, Ortiz, López, Núñez, Carrasco y Bilbao (2019); PNCE (2019).	Perales, Arias & Bazdresch (2014).	Ley de Violencia Escolar
Gestión de la convivencia	Entrega de apoyo a los establecimientos educacionales. Se intenta transmitir valores asociados a la buena convivencia, enfoque proactivo, dialógico y de responsabilidad compartida. Estrategias escolares de tipo formativas, normas construidas de forma participativa.	Enfoque preventivo y centrado en el desarrollo humano, la educación socioemocional y el autocuidado de los estudiantes y docentes. Busca gestionar las emociones para autorregular las conductas y propiciar que las emociones se manifiesten constructivamente. Estrategias de educación emocional para conocer y manejar las emociones propias (inteligencia emocional) y reconocer las emociones de los demás, lo que permite adecuarse a esas emociones (inteligencia interpersonal).	Enfoque centrado en la construcción de espacios seguros y de evitación de la violencia e indisciplina, reactivo frente a la violencia, con una relación jerárquico-autoritaria entre alumno y profesor, basada en la autoridad de este sobre las y los alumnos. Recurre a estrategias punitivas basadas en la aplicación de sanciones, como suspensiones y expulsión de “individuos conflictivos”, que son considerados como la causa de los conflictos.
Valores de la gestión de la convivencia	Participación, transformación social, inclusión, justicia, paz, solidaridad, colaboración, igualdad, deliberación, pluralismo.	Tolerancia, respeto mutuo, amabilidad, confianza, autoconocimiento, cuidado de sí, de los otros y del mundo que les rodea.	Disciplina, seguridad, orden, tolerancia cero a la violencia y la indisciplina, armonía en el grupo, respeto a la autoridad.
Fines de la gestión de la convivencia	Construir sociedades más democráticas desde las escuelas.	Desarrollar habilidades socioemocionales en los actores de las comunidades educativas.	Reducir la violencia en las escuelas para favorecer aprendizajes disciplinares.

Fuente: Elaboración Propia en base a la literatura consultada.

2.2 ¿Qué factores influyen en la convivencia escolar?

En este punto se exponen los factores que influyen en la convivencia escolar entendida “*como preparación para la democracia*”, de acuerdo con la clasificación del punto 2.1, y siguiendo el punto de vista del Programa PACES-PUCV.

Según la revisión bibliográfica de este estudio existen factores de riesgo y factores protectores para una buena convivencia en los establecimientos educacionales, que pueden agruparse en: **a)** factores individuales; **b)** factores familiares; **c)** factores escolares; **d)** factores contextuales o

asociados a las comunidades o barrios y e) factores estructurales. El supuesto del modelo ecológico es “*entender que los comportamientos individuales no se explican únicamente por factores de orden individual, pues éstos están insertos y son interdependientes de contextos*” (Ascorra y López, 2019: 28). De esta forma, los factores asociados a la convivencia en la escuela no son únicos sino múltiples, con una expresión local específica, por lo que un diagnóstico debiese considerar siempre el caso concreto del establecimiento de forma contextualizada.

Además, algunos factores serán más susceptibles de ser abordados por unos actores que por otros. Los factores escolares, por las comunidades educativas, mientras que los de política pública serán abordables en menor grado por las comunidades escolares y en mayor grado por actores extraescolares, como autoridades políticas. Por último, es importante resaltar que una clasificación de este tipo resulta de cierta forma arbitraria debido al entrelazamiento de factores que ocurre en la práctica, aunque se considera que tiene valor expositivo y didáctico.

2.2.1 Factores individuales

Los factores individuales son los que dependen de la persona y de sus características particulares, incluyendo su género, sexo, la etapa en que se encuentran en el ciclo vital individual, características psicológicas, características sociales, modelo de creencias, etnicidad, entre otros.

Entre los factores de riesgo individuales para una buena convivencia escolar se encuentran la impulsividad, bajo nivel educativo, el abuso de sustancias psicotrópicas y antecedentes de comportamiento agresivo o de haber sufrido maltrato (Tuvilla, 2004).

Por su parte, como factores de protección, las habilidades socioemocionales de los integrantes de la comunidad educativa juegan un rol significativo. Tuvilla (2004:18) enumera como factores de protección: “*una percepción positiva de sí mismo a través del desarrollo de la autoestima, la adquisición de aquellas competencias sociales como la capacidad de pedir ayuda y reaccionar con energía antes las diferentes manifestaciones de violencia; alcanzar un nivel de concientización adecuado respecto a diferentes formas o expresiones de la violencia; y adquirir habilidades y aptitudes que favorezcan el apoyo entre víctimas y agresores*”. Esta enumeración de factores es válida para todos los miembros de la comunidad escolar, pues su autorregulación, empatía, colaboración y desarrollo ético juegan un rol en la calidad de la convivencia escolar.

La escuela enseña modos de convivir, pero cada uno convive de acuerdo con modos aprendidos en distintos espacios sociales, sean estos del signo que sean.

Por último, es importante señalar que existen estudios que han evidenciado que el foco en las características individuales del estudiantado es un foco limitado cuando hablamos de intervenciones integrales para la mejora de la convivencia escolar (Ascorra y López, 2019).

2.2.2 Factores familiares

Los factores familiares son aquellos que dependen del entorno social más inmediato del individuo, incluyendo la cohesión familiar, el estilo educativo familiar, entre otros.

Como señala Tuvilla (2004), las relaciones sociales tienen el potencial de configurar el comportamiento de un individuo. En particular, las relaciones familiares tienen un impacto considerable para aprender a convivir con otros, aunque no se debe olvidar que, en cuanto relaciones, las familiares son también moldeables. Las personas tienen capacidad de agencia en sus relaciones, aunque en ciertos contextos donde las vulnerabilidades se potencian entre sí, esta capacidad parece debilitarse.

Llevado a los centros educativos, *“los y las estudiantes pueden incidir, también, en las pautas de interacción de sus propias familias, a partir de lo que observan y aprenden en la escuela”* (López, Díaz & Carrasco, 2015:9) aunque ciertos entornos familiares son el lugar donde los jóvenes aprenden la violencia (Tuvilla, 2004). Es así que Fernández (1999) citada en Martínez-Otero (2001) expone como factores que pueden aumentar la agresividad en niños y adolescentes la desestructuración familiar, malos tratos y utilización de violencia, la falta de diálogo, y permisividad, indiferencia o excesiva punición.

2.2.3 Factores escolares

Los factores escolares son los que dependen de la escuela y de su capacidad para gestionar los resultados académicos y no académicos de los estudiantes. Estos factores incluyen desde intangibles como la cultura de la escuela hasta el nivel de apoyo a los estudiantes, sus niveles de victimización y la misma infraestructura que posee el establecimiento.

El rol de las creencias también es importante. El que algunos estudiantes vean la escuela como un lugar de aparcamiento forzoso más que como una vía de realización posterior, genera tensiones sin fin (Subirats & Alegre, 2006). Los modos de convivir en la escuela pueden también configurarse en función de lo que Subirats y Alegre denominan “culturas adolescentes”, que pueden estar en oposición a la “cultura escolar”, como los modos de pensar adultos pueden estarlo de los modos de pensar de los estudiantes. En ocasiones esas culturas coexisten con la cultura y reglas escolares mientras que en otras circunstancias estas culturas actúan en la escuela transgrediendo, en mayor o menor medida, sus marcos normativos. Pero esta contestación a las normas e indisciplina no siempre significa lo mismo, ni tiene igual fuerza e intensidad. En este punto los autores distinguen entre la “resistencia” y el “gamberrismo”. La resistencia incluye el descrédito de los valores instrumentales de la escuela, pensar la escolarización como un proceso engañoso y sentir escepticismo respecto a sus promesas de preparación, inserción laboral, movilidad social, etc. El gamberrismo, correspondería por su parte, a actos de provocación producto del aburrimiento, el cansancio, la desidia y la búsqueda de distinción dentro del grupo por medio de la demostración de sus capacidades de atrevimiento (Subirats y Alegre, 2006).

En cuanto a la “cultura escolar”, entendida como *“los conocimientos socialmente compartidos y transmitidos sobre lo que es y lo que debería ser, que se transmiten de manera voluntaria e implícita, y se simbolizan a través de actos y productos”* (Gather, 2004) citado en Banz (2008:2), es, de cierta forma, el recipiente de la convivencia. La cultura escolar siempre plantea formas de convivencia que están asociadas a alguna funcionalidad para los actores y su tiempo (Banz,

2008), y es un espacio no físico en que las comunidades escolares habitan, donde se moldean las actitudes, valores y creencias que influyen en el comportamiento de cada actor de una comunidad.

Y aunque las maneras de convivir terminan asumiéndose como “lo que es” y “debería ser” (Banz, 2008) se trata de una construcción colectiva, por lo que las personas tienen la posibilidad de gestionarla y transformarla en direcciones correctas o incorrectas. En esta línea, Martínez-Otero (2001) señala como factores que pueden desencadenar violencia en las escuelas, entre algunos otros, la crisis y diversidad de valores en la institución escolar y la presencia de una cultura escolar hegemónica que puede chocar con otra en posición desventajosa.

En el año 2015 se publicó el libro “¡Nosotros sí podemos!: Aprendiendo a mejorar la convivencia escolar”, que recogió experiencias de escuelas con buena calidad de convivencia escolar, en que la dimensión cultural es trabajada con énfasis especial. Para estas escuelas trabajar con la convivencia escolar está considerado dentro del proyecto educativo, se gestiona activamente, considera valores, de forma inter-estamental, con rituales y tradiciones que consideran el proyecto educativo-valórico (López, Díaz y Carrasco, 2015). En uno de los artículos finales del libro se destaca la razón de ser de una buena práctica “*que conlleva la obtención de resultados y construcción de una cultura escolar en función de la formación integral y aprendizajes de los y las estudiantes*” (López, Díaz y Carrasco, 2015:116). Las buenas prácticas reflejan elementos que permiten adentrarse en la cultura de la escuela y los significados que ésta conlleva destacándose como elemento común querer contribuir a la formación integral y los aprendizajes de todos los estudiantes (López, Díaz y Carrasco, 2015).

Además, un clima escolar positivo es un factor protector de la convivencia. Ma y Huebner (2008) citados por Ascorra y López (2019) evidencian que los buenos climas escolares aumentarían la satisfacción y percepción de cercanía con los pares. En contraste, un mal clima escolar impacta en la satisfacción con la vida de los estudiantes, la cual disminuye a mayor violencia, física o psicológica, recibida por el grupo (Bazán & Oyanedel, 2011 citado por Ascorra y López, 2019).

2.2.4 Factores contextuales

Estos factores se refieren al ambiente social y físico en que se inserta la escuela. Incluye las características del barrio (nivel de pobreza, deterioro físico, apoyo institucional) hasta problemas sociales como el consumo y tráfico de drogas, delincuencia, entre otros.

Estudios también han indagado la relación entre segregación y convivencia escolar en escuelas públicas chilenas, evidenciando “*peores ambientes escolares en las escuelas municipales, de niveles socioeconómicos bajos y no selectivos, donde los estudiantes se sienten estigmatizados y discriminados por variables académicas, familiares y políticas*” (Ascorra, López, Núñez, Bilbao, Gómez y Morales, 2016).

2.2.5 Factores de política pública

Estos factores se refieren al entorno institucional en que las escuelas desarrollan su labor, entendiendo por “institución” las “reglas del juego” en que ellas desenvuelven su actividad.

Respecto a esas reglas, en Chile se ha hablado de un ambiente punitivo en la política de convivencia escolar, ya que el país ha tenido cuatro políticas distintas de convivencia escolar (2004, 2011, 2015 y 2019) y una de violencia escolar (2011) evidenciándose tanto una lógica punitiva como formativa (Ascorra y López, 2019). Debido al peso de las leyes promulgadas este clima se ha vuelto progresivamente punitivo. Es así que “*la Ley SEP y la Ley de Violencia Escolar, obligan a la escuela a mejorar la convivencia escolar, amenazándolas con medidas de castigo, menos recursos, menos autonomía, y, finalmente, con su cierre en caso de no resultar competentes*” (Ascorra & López, 2019:163). Reforzando esta idea, las autoras señalan:

“La gestión de la convivencia escolar se complejiza aún más, cuando evidenciamos que los objetivos que persigue la política pública van variando en el tiempo. Morales y López, (2019) identificaron una política de convivencia escolar ambivalente, que va de lo punitivo a lo formativo; de lo democrático, a la seguridad escolar, pasando por enfoques de salud mental restaurativa y perspectivas managerialista. Lo anterior pone de relieve que las finalidades últimas y la forma de gestionar la convivencia en las escuelas chilenas no son tan claras como se podría esperar” (2019: 141).

Estas políticas desempeñan un rol importante en modelar el comportamiento de los actores del sistema educativo, pero, en la medida que apuntan a objetivos diferentes, tienen un efecto restringido en la mejora de la convivencia escolar. En este sentido, las políticas de convivencia escolar en el país no corresponden a un conjunto estructurado de estrategias, sino a un grupo de medidas, más o menos aisladas unas de otras, con distinta importancia relativa para los actores de las comunidades educativas de acuerdo con sus respectivos roles. En la Tabla N° 2 se resumen los factores expuestos previamente:

Tabla N°2. Resumen de factores que influyen la convivencia escolar.

Tipología	Factores
Factores individuales	<ul style="list-style-type: none"> • Habilidades socioemocionales de cada integrante de la comunidad educativa. • Características físicas y psicológicas de cada integrante de la escuela. • Factores sociodemográficos: clase, género, etnia, edad. • Condición médica de cada integrante de la comunidad educativa. • Comportamientos problemáticos. • Dificultades de aprendizaje.
Factores familiares	<ul style="list-style-type: none"> • Condiciones laborales de las familias. • Nivel educativo de los padres y capital cultural de su entorno. • Nivel socioeconómico de las familias. • Nivel de involucramiento de los padres con la escuela. • Tipo de convivencia familiar (violencia intrafamiliar, abusos a niños, conflictos maritales) como escuela de resolución de conflictos.
Factores escolares	<ul style="list-style-type: none"> • Cultura de la escuela (creencias, valores, etc.) y culturas adolescentes. • Características de establecimiento (dependencia administrativa, ubicación, desempeño académico y no académico, etc.) • Satisfacción o desafección con la escuela. • Apoyo al estudiante • Participación de los estudiantes en actividades escolares.

Tipología	Factores
	<ul style="list-style-type: none"> • Victimización verbal, física, acoso sexual, de profesores, a profesores, entre profesores. • Presencia de robos, deterioro premeditado del material educativo, conformación de grupos y pandillas rivales, presencia de armas. • Ambiente físico, organización y participación en el curso, relación entre pares, orientación al aprendizaje.
Factores contextuales	<ul style="list-style-type: none"> • Características del barrio (nivel de pobreza, deterioro físico, apoyo institucional). • Presencia de narcotráfico y drogadicción en el barrio. • Índices de delincuencia, desigualdad entre territorios, segregación socioespacial. • Influencia de los medios de comunicación
Factores de política pública	<ul style="list-style-type: none"> • Políticas públicas educativas. • Políticas sociales en general (vivienda, seguridad ciudadana, otras)

Fuente: Elaboración propia en base a literatura consultada.

2.3 Medidas de la convivencia escolar en Chile

2.3.1 Indicador Clima de Convivencia Escolar

El IDPS ICCE considera las percepciones de estudiantes, docentes y padres y apoderados sobre la presencia de un ambiente de respeto, organizado y seguro en la escuela (Mineduc, 2019b). El ambiente de respeto considera las percepciones y actitudes en relación con el trato respetuoso entre los miembros de la comunidad educativa, la valoración de la diversidad y la ausencia de discriminación. Un ambiente organizado considera las percepciones y actitudes sobre la existencia de normas claras, conocidas exigidas y respetadas por todos y el predominio de mecanismos constructivos de resolución de conflictos. Por último, el ambiente seguro alude a las percepciones y actitudes con relación al grado de seguridad y de violencia física y psicológica al interior de la escuela, así como la existencia de mecanismos de prevención ante ésta (ACE, 2016). El documento “Informe técnico 2017: Indicadores de Desarrollo Personal y Social (IDPS) medidos a través de cuestionarios” contiene una explicación más detallada de su construcción.

El ICCE intenta medir censalmente de qué forma las comunidades educativas experimentan el clima de convivencia escolar. Este indicador se construye en base a la información recopilada por los Cuestionarios de Calidad y Contexto de la Educación, que corresponden a un conjunto de instrumentos que evalúan percepciones, actitudes y conductas de los actores de las comunidades educativas sobre aspectos del ámbito escolar y familiar de los estudiantes (ACE, 2017) y que son aplicados a estudiantes, docentes, padres y apoderados junto al SIMCE.

La Tabla N° 3 muestra las preguntas e ítems usados para medir la presencia en un ambiente de respeto, organizado y seguro el año 2016 para estudiantes de 6° año de educación básica:

Tabla N° 3. Encuesta IDPS Clima de Convivencia Escolar 6° básico 2016.

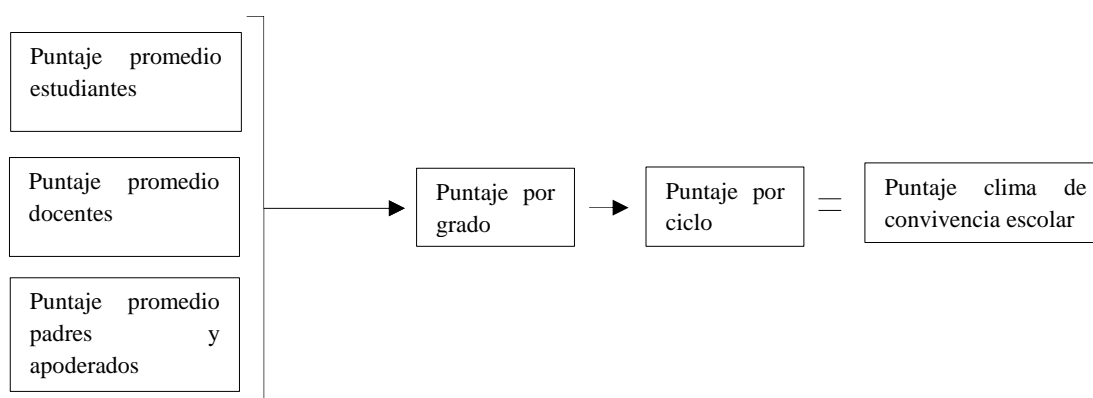
N°	Pregunta	Ítem
1	¿Cuántas veces pasa lo siguiente en tu curso?	Los profesores de mi curso nos tratan con respeto.

N°	Pregunta	Ítem
2	¿Cuántas veces pasa lo siguiente en tu curso?	Los estudiantes de mi curso mantenemos limpia la sala de clases.
3	¿Cuántas veces pasa lo siguiente en tu curso?	Los estudiantes de mi curso nos tratamos con respeto.
4	¿Cuántas veces pasa lo siguiente en tu escuela?	Los estudiantes de mi escuela se tratan bien.
5	¿Cuántas veces pasa lo siguiente en tu escuela?	Los estudiantes cuidan que la escuela esté limpia (por ejemplo, los baños, patios, cancha, etc.).
6	¿Tus profesores han enseñado lo siguiente a tu curso este año?	Mis profesores nos han enseñado que hay que respetar a todas las personas, aunque sean diferentes.
7	¿Cuántas veces pasa lo siguiente en tu curso?	Los estudiantes de mi curso trabajamos siguiendo las instrucciones que da el profesor.
8	¿Cuántas veces los profesores de tu curso hacen lo siguiente en clases?	Mis profesores explican las consecuencias que tiene para mí o mis compañeros hacer algo incorrecto.
9	¿Cuántas veces los profesores de tu curso hacen lo siguiente en clases?	Mis profesores enseñan la importancia de hacer las tareas y trabajos a tiempo.
10	¿Cuántas veces los profesores de tu curso hacen lo siguiente en clases?	Mis profesores explican por qué es importante cumplir las normas.
11	¿Qué te parecen las siguientes acciones?	Copiar en las pruebas.
12	¿Qué te parecen las siguientes acciones?	Quitarle la colación a un compañero.
13	¿Qué te parecen las siguientes acciones?	Dañar cosas del establecimiento (por ejemplo, romper o rayar mesas, sillas, vidrios, computadores, etc.).
14	¿Tus profesores han enseñado lo siguiente a tu curso este año?	Mis profesores nos han enseñado qué hacer si un compañero nos pega.
15	¿Tus profesores han enseñado lo siguiente a tu curso este año?	Mis profesores nos han enseñado qué hacer si un compañero le pega a otro compañero.
16	¿Tus profesores han enseñado lo siguiente a tu curso este año?	Mis profesores nos han enseñado qué hacer si un compañero nos quita los materiales.
17	¿Tus profesores han enseñado lo siguiente a tu curso este año?	Mis profesores nos han enseñado qué hacer cuando a un compañero lo molestan mucho.
18	¿Tus profesores han enseñado lo siguiente a tu curso este año?	Mis profesores nos han enseñado que está mal robar.
19	¿Tus profesores han enseñado lo siguiente a tu curso este año?	Mis profesores nos han enseñado que está mal pegarles a los compañeros.
20	¿Cuántas veces durante este año te han pasado estas cosas en tu curso?	Algún compañero se ha burlado de ti.
21	¿Cuántas veces durante este año te han pasado estas cosas en tu curso?	Algún compañero te ha pegado.
22	¿Cuántas veces durante este año ha pasado lo siguiente en tu escuela?	Tú o alguno de tus amigos han sentido miedo de que otros estudiantes los empujen, les peguen o los encierren dentro de la escuela.
23	¿Cuántas veces durante este año ha pasado lo siguiente en tu escuela?	Algunos estudiantes han obligado a hacer cosas que no querían a otros compañeros.
24	¿Cuántas veces durante este año ha pasado lo siguiente en tu escuela?	Algunos estudiantes de la escuela han amenazado con pegarles a otros compañeros.

Fuente: Agencia de Calidad de la Educación.

El ICCE estima en forma independiente los resultados para la educación básica y la educación media (Mineduc, 2019b). El puntaje por ciclo (básica o media) se calcula en base a todos los grados correspondientes, para lo que se promedian los puntajes de cada grado y se ponderan por el número de estudiantes que rindió el SIMCE en cada uno de ellos. Para calcular el puntaje por grado se agregan los puntajes promedio de estudiantes, docentes y padres y apoderados que responden el cuestionario, asignándoles una ponderación de 50% al puntaje de estudiantes, 10% a docentes y 40% a padres y apoderados. Por último, el puntaje de cada estudiante, docente y padre se basa en su respuesta a las preguntas del cuestionario relacionadas con el indicador (Mineduc, 2019b). Esquemáticamente el proceso de cálculo se muestra en la Figura N°1:

Figura N°1. Proceso de elaboración del ICCE.



Fuente: Mineduc 2019b.

2.3.2 ¿Un indicador desde el modelo de la Encuesta PACES-PUCV?

El Programa PACES-PUCV ha desarrollado un Sistema de Monitoreo de la Convivencia Escolar inspirado en el modelo social-ecológico, que busca poner a disposición de las comunidades educativas información clara, precisa y relevante para la gestión de la convivencia escolar. Entre las herramientas que se el Programa PACES-PUCV proporciona se encuentra el “Subsistema de Medición y Evaluación de la Convivencia Escolar” (López et al., 2013), que permite enriquecer la mirada sobre la convivencia en las escuelas y facilita al usuario final su interpretación y utilización. De esta manera, *“el subsistema de medición y evaluación de la convivencia escolar se presenta físicamente como una batería de medición compuesta por tres encuestas: una para ser aplicada a estudiantes, otra para los profesores y otra para los apoderados. Las condiciones de uso implican su aplicación a través de un sistema on line de aplicación que es presencial en las escuelas. Este sistema de medición y evaluación permite medir 9 escalas no evaluadas masivamente antes en el sistema escolar, entregando información pertinente y adecuada a las escuelas para la toma de decisiones”* (López et al., 2013: 206).

En el mismo artículo las autoras señalan brevemente cuáles son las diferencias entre las encuestas aplicadas a estudiantes, profesores y padres, madres y/ apoderados. Informan que la encuesta aplicada a estudiantes “*incluye nueve dimensiones: clima escolar, victimización de estudiantes, victimización de adultos a estudiantes, intimidación de profesores, porte de cuchillo, peligrosidad en la escuela, peligrosidad en el barrio, contribución a la escuela y satisfacción con la vida*” (López et al., 2013: 206). Por su parte, “*la encuesta a profesores incluye 5 dimensiones: conductas de riesgo observadas en los estudiantes; victimización a profesores; percepción de seguridad personal en la escuela; capacitación y formación continua en convivencia escolar; prevención y respuestas a la convivencia escolar*” (López et al., 2013: 206). Finalmente, la encuesta a padres, madres y/o apoderados de estudiantes “*incluye 5 dimensiones: satisfacción con la escuela; participación; normas antibullying; gestión de la convivencia y victimización a apoderados*” (López et al., 2013: 207).

Así, el Programa PACES-PUCV desarrolló una herramienta de medición que incluye múltiples dimensiones de la convivencia, muchas de las cuales no están incluidas en el ICCE de la Agencia de Calidad de la Educación, desde el punto de vista de cada actor de la comunidad escolar, por separado. Además, como indican las autoras, “*las propiedades psicométricas del subsistema de medición, presentadas aquí, nos permiten concluir que la encuesta para estudiantes es fiable y válida*” (López et al., 2013: 216). Señalando, además, que “*las limitaciones de este trabajo no se encuentran en el instrumento, el cual ha quedado validado como un instrumento fiable y válido, sino más bien en el proceso mismo de monitoreo*” (López et al., 2013: 217).

El instrumento de PACES para alumnos utiliza las escalas de “clima escolar”, “clima de aula”, “victimización” y “Seguridad en la escuela”. Dentro de esas escalas se incluyen quince subdimensiones, mostrando la complejidad de la medición de la convivencia escolar. Esta complejidad se expresa también en las noventa preguntas que desglosan distintas partes del fenómeno convivencial. En el siguiente cuadro se muestra una comparación general entre estas formas de comprender y medir la convivencia escolar:

Tabla N°4. Cuadro comparativo de mediciones de convivencia escolar.

Ámbito	Indicador ICCE	Encuesta PACES
Fuentes de información	Cuestionarios de Calidad y Contexto de la Educación.	Encuesta PACES a estudiantes.
Características de la encuesta	Preguntas cerradas, por lo generales con escalas de cuatro alternativas (escala de Likert).	Preguntas cerradas, por lo generales con escalas de cuatro alternativas (escala de Likert).
Diferencias entre las encuestas	24 preguntas en tres dimensiones	Más de 80 preguntas en cuatro dimensiones
Dimensiones consideradas	<ol style="list-style-type: none"> 1. Ambiente de respeto 2. Ambiente organizado 3. Ambiente seguro en la escuela 	<ol style="list-style-type: none"> 1. Clima escolar 2. Clima de aula 3. Victimización entre pares 4. Seguridad en la escuela
Metodología de Cálculo	Promedio ponderado de los puntajes de padres (40%), docentes (10%) y estudiantes (50%), obteniendo un puntaje	Elaboración propia en base a los datos del Programa PACES-PUCV.

Ámbito	Indicador ICCE	Encuesta PACES
	por grado. Luego realiza un promedio simple para obtener el puntaje por ciclo y el indicador para el establecimiento.	
Alcance	Nacional	Local

Fuente: Elaboración propia.

Ambas medidas tienen en común que miden la convivencia escolar de forma descriptiva y retrospectiva, y responden a qué sucedió en el pasado con la convivencia en las escuelas.

2.4 Predicción y aprendizaje automático

En términos genéricos el aprendizaje automático o “*machine learning*” es un área en auge que consiste en una serie de métodos, modelos y algoritmos utilizados para el análisis de datos, con el fin de mejorar la toma de decisiones en todas sus etapas. Esta disciplina permite realizar predicciones a partir de los datos disponibles y se centra en buscar patrones en ellos para elaborar predicciones sobre lo que pasará en el futuro. El aprendizaje automático ha sido definido como una subespecialidad de la Ciencia de la Computación que se ocupa del diseño y desarrollo de algoritmos que permiten inferir comportamientos en base a datos empíricos (Rodríguez, Palomino & Mondaca, 2017) o la aplicación y ciencia de los algoritmos que da sentido a los datos, para convertir estos datos en conocimiento (Raschka & Mirjalili, 2019), a través de los métodos y algoritmos utilizados o un conjunto de métodos que pueden detectar automáticamente patrones en los datos y luego usar los patrones descubiertos para predecir el futuros datos o para realizar otros tipos de toma de decisiones en condiciones de incertidumbre (Murphy, 2012).

Las aplicaciones del aprendizaje automático se han documentado en el ámbito público y privado, en planos tan distintos como determinar la ubicación de una persona, la amistad, el amor, compras, ventas y cancelaciones realizadas, correos spam, lesiones en el puesto de trabajo, reclamaciones a las compañías de seguro, saldos impagos, nacimientos prematuros, fraudes, delincuencia callejera, reincidencia delictual, tráfico vehicular, educación, entre muchos otros temas (Siegel, 2013). A su vez, Rodríguez, Palomino & Mondaca (2017) señalan que se han generado aplicaciones de “*machine learning*” para estimar automáticamente el nivel socioeconómico de una cierta zona geográfica en base a información satelital con el fin de establecer políticas de ayuda social, otras para estimar el riesgo de deserción escolar de un alumno, escogiendo qué intervenciones son más costo-efectivas y mejorar las políticas de fiscalización usando inspecciones predictivas basadas en reseñas en línea de clientes.

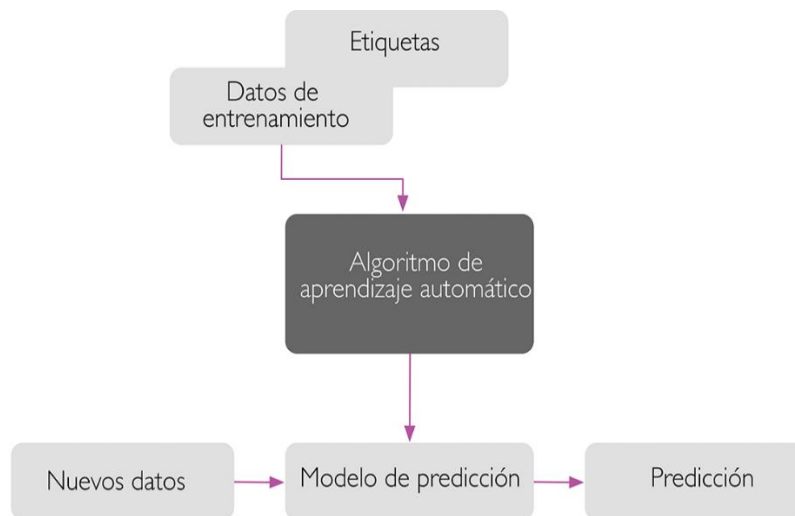
Aunque actualmente no es posible predecir con precisión, las predicciones no tienen por qué ser precisas para ser valiosas y útiles (Siegel, 2013). Además, es importante tener presente que una correlación no implica causalidad y que el descubrimiento de una relación predictiva entre A y B no significa que una cause a la otra ni siquiera indirectamente. Es en ese sentido que la analítica predictiva no busca tanto la causalidad como la funcionalidad, ya que está orientada a la acción (Siegel, 2013). Estos problemas no necesitan establecer una causalidad para apoyar la toma de

decisiones informada, en el escenario típico de evaluación con o sin la implementación de una política pública. “Por ejemplo: ¿es la posibilidad de lluvia lo suficientemente alta para salir con paraguas?: en dicho caso, no se necesita saber qué causa la lluvia, solo estimar si se producirá o no” (Rodríguez, Palomino & Mondaca, 2017: 22)

Para realizar su tarea de predecir, el “*machine learning*” dispone de tres tipos de aprendizajes cuyas diferencias fundamentales se exponen a continuación:

- a) **Aprendizaje supervisado:** trabaja aprendiendo de un conjunto de datos de entrenamiento que están “etiquetados”. Luego de entrenar a los algoritmos con datos etiquetados se prueba la capacidad predictiva del algoritmo con un conjunto de datos nuevos para los cuales no se tiene etiqueta. El término “supervisado” se refiere a un conjunto de muestras donde las señales de salida deseadas (etiquetas), ya se conocen (Raschka & Mirjalili, 2019). Este tipo de aprendizaje se usa con frecuencia en problemas de clasificación. Para valores continuos utiliza la regresión mientras que para predecir valores discretos podemos usar K-nearest neighbors (KNN), XGboost o Catboost. La siguiente figura ilustra el funcionamiento de esta clase de aprendizaje:

Figura N°2 Funcionamiento del aprendizaje supervisado.



Fuente: Raschka & Mirjalili, 2019.

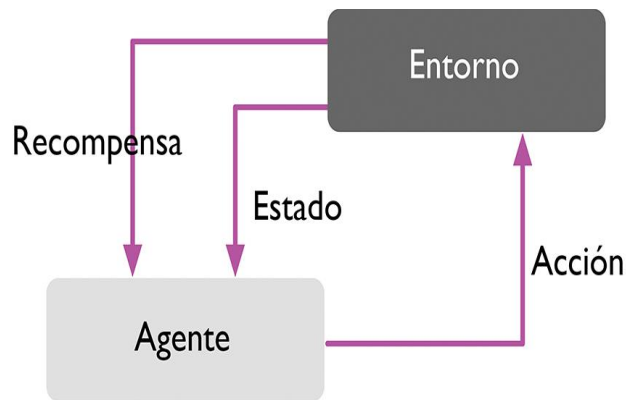
- b) **Aprendizaje no supervisado:** a diferencia del aprendizaje supervisado, este tipo de aprendizaje no cuenta con datos “etiquetados” para entrenar sus algoritmos y tiene como objetivo encontrar algún patrón en los datos que analiza pues su estructura es desconocida. Con las técnicas de aprendizaje sin supervisión podemos explorar la estructura de nuestros datos para extraer información significativa sin la ayuda de una variable de resultado conocida o una función de recompensa (Raschka & Mirjalili, 2019).

Ejemplos de tipos de algoritmos de aprendizaje no supervisado son los que permiten construir agrupaciones (como K-means o K-medias en español) y los que permiten descubrir factores latentes (reducción de dimensionalidad) (Murphy, 2012).

- c) **Aprendizaje reforzado:** este tipo de aprendizaje se basa en modificar la respuesta del algoritmo utilizando retroalimentación (recompensas y castigos). En este tipo de aprendizaje el objetivo es desarrollar un sistema (agente) que mejore su rendimiento basado en las interacciones con el entorno (Raschka & Mirjalili, 2019). Con estas técnicas se busca simular el proceso de aprendizaje humano. Por ejemplo, en ajedrez, el resultado de cada movimiento podría ser un estado distinto del entorno, asociadas a un evento positivo (como eliminar la pieza del contrincante) o negativo (como perder una pieza en una jugada). Pero no todos los turnos dan como resultado la eliminación de una pieza por lo que el aprendizaje reforzado se centra en aprender series de pasos maximizando una recompensa en base a retroalimentación inmediata y diferida (Raschka & Mirjalili, 2019).

La siguiente figura ilustra el funcionamiento de esta clase de aprendizaje:

Figura N°3. Funcionamiento del aprendizaje reforzado.



Fuente: Raschka & Mirjalili, 2019.

Para resumir, las características generales de estos tres de tipos de aprendizaje automático se muestran brevemente en la Figura N°4:

Figura N° 4. Tipos de Aprendizaje Automático.

Aprendizaje supervisado	<ul style="list-style-type: none">> Datos etiquetados> <i>Feedback</i> directo> Predicción de resultados/futuro
Aprendizaje no supervisado	<ul style="list-style-type: none">> Sin etiquetas> Sin <i>feedback</i>> Encontrar estructuras ocultas en los datos
Aprendizaje reforzado	<ul style="list-style-type: none">> Proceso de decisión> Sistema de recompensa> Aprender series de acciones

Fuente: Raschka & Mirjalili, 2019.

Luego de conocer en general los tipos de aprendizaje automático debemos identificar la clase de problema de predicción que se busca resolver. Dependiendo del problema debemos seleccionar las herramientas adecuadas, se trate de una tarea de: a) clasificación; b) regresión; c) agrupamiento o d) reducción dimensional. Si queremos predecir una clase, el problema será de clasificación; si nuestro objetivo es predecir un valor numérico continuo, el problema es de regresión; si queremos agrupar los datos tenemos un problema de agrupamiento (clustering) y, por último, si queremos reducir la cantidad de variables manteniendo la riqueza de la información original, se trata de un problema de reducción de dimensionalidad⁶.

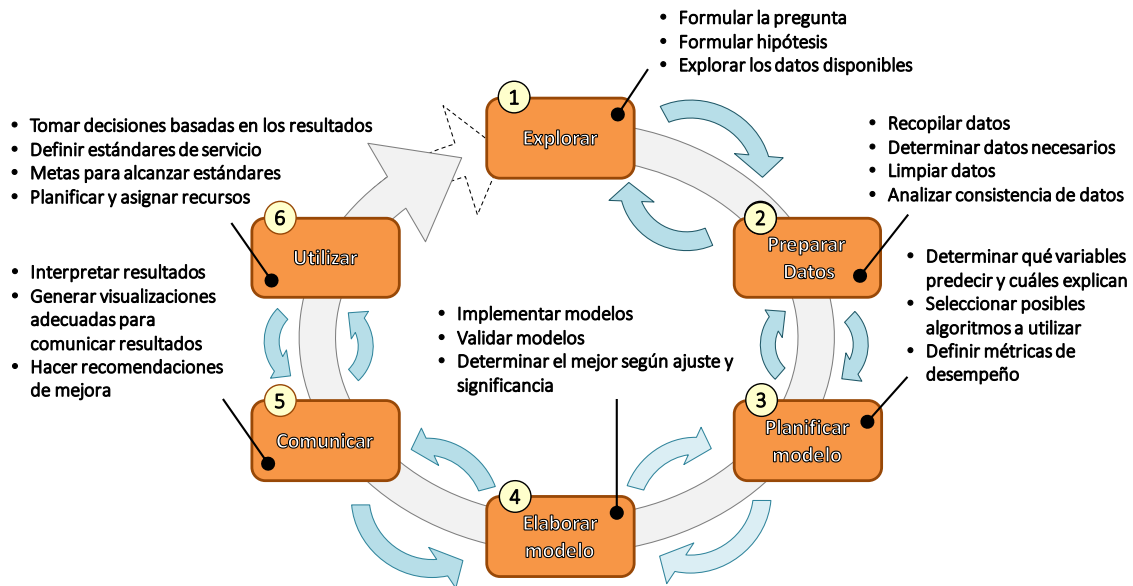
En este caso lo que se busca predecir es el resultado de la medición de la Encuesta PACES, es decir: a) un único resultado numérico que nos indica la situación en que se encuentra la convivencia en las escuelas, lo que significaría usar herramientas de regresión o bien b) un valor discreto que clasifique a las escuelas en su tipo de convivencia escolar “buena” o “débil”, lo que nos enfrentaría a un problema de clasificación propio del aprendizaje supervisado.

⁶ La siguiente figura de scikit-learn nos muestra los tipos de algoritmos a utilizar dependiendo de las características del problema que se busque resolver: <https://bit.ly/3mmT7XY>

3 Marco Metodológico

En este capítulo se describen la metodología y herramientas para conseguir los objetivos del estudio. Esta investigación tiene un carácter cuantitativo y predictivo, en la medida que busca hallar una serie de variables que predigan la evolución de fenómenos concretos (Cea, 2001). Para comparar la capacidad de distintos modelos de aprendizaje supervisado de predecir el resultado de la Encuesta PACES, se siguieron los pasos del análisis de datos: 1) Explorar datos, 2) Preparar datos, 3) Planificar un modelo, 4) Elaborar un modelo, 5) Comunicar resultados y 6) Utilizar los hallazgos para la toma de decisiones (Schmarzo, 2013). A continuación, se muestra el esquema del proceso en la Figura N°5:

Figura N°5. Ciclo de vida de análisis de datos.



Schmarzo (2012): Data Analytics Lifecycle

Fuente: Schmarzo, 2012.

Este proceso se describe brevemente a continuación:

- 1) La etapa de **Explorar datos** consiste en obtener una comprensión detallada de los procesos y el dominio del negocio. Incluye capturar las preguntas y decisiones de negocio más importantes, así como evaluar los recursos disponibles (Schmarzo, 2013).
- 2) La etapa de **Preparar datos** consiste en determinar los datos necesarios para el problema de investigación, adquirirlos, limpiarlos y analizarlos. Esto incluye el uso de técnicas de visualización de datos y herramientas para comprenderlos, identificando (y eliminando si es necesario) valores atípicos y evaluando lagunas para determinar la calidad general de los datos (Schmarzo, 2013). En esta etapa es importante un análisis exploratorio que nos permita vislumbrar posibles patrones y reconocer distribuciones estadísticas útiles en el futuro, a la luz del objetivo del trabajo. También es importante identificar cuántos datos

tenemos, la existencia de valores nulos, cuáles son discretos y continuos, los tipos de datos disponibles (texto, entero, decimales, etc.), si es problema es de tipo supervisado o no supervisado, cuál es la columna binaria o multiclase que buscamos predecir, si está o no balanceada la base de datos, cuáles son las variables más importantes para nuestro problema, hay correlación entre ellas, entre otras (Bagnato, 2020).

- 3) La etapa de **Planificar el modelo** incluye las siguientes actividades: a) determinar los diferentes modelos, métodos, técnicas y flujos de trabajo analíticos a explorar como parte del desarrollo del modelo analítico; b) determinar la correlación y colinealidad entre variables para seleccionar variables clave que se utilizarán en el desarrollo del modelo (Schmarzo, 2013).
- 4) La etapa de **Elaborar el modelo** incluye las siguientes actividades: a) utilizar los conjuntos de datos para pruebas, entrenamiento y producción; b) evaluar la viabilidad y confiabilidad de los datos para usar en los modelos predictivos. y c) finalmente, desarrollar, probar y perfeccionar los modelos analíticos (Schmarzo, 2013). Para Schmarzo (2013) esta es una etapa altamente iterativa, en que se fallará varias veces al probar diferentes variables y técnicas de modelado.
- 5) En la etapa de **Comunicar Resultados** se incluyen las siguientes actividades: a) determinar la calidad y confiabilidad del modelo analítico y la significancia estadística, mensurabilidad y capacidad de acción de los conocimientos analíticos resultantes y b) desarrollar los cuadros y gráficos para comunicar los conocimientos, los resultados y las recomendaciones del modelo analítico (Schmarzo, 2013).
- 6) En esta última etapa de **Utilizar** se incluyen las siguientes actividades: a) entrega de las recomendaciones finales, informes, resúmenes, código y documentos técnicos; b) implementar los modelos analíticos en los entornos productivos y operativos. (Schmarzo, 2013). Si bien el autor considera otras actividades dentro de esta etapa, no se consideran pertinentes por tratarse de una tesis de grado.

Una vez generado el modelo se imputará el resultado de la medición de la de convivencia escolar de PACES a escuelas del país en base a datos administrativos. Para comprobar su efectividad se le pidió al modelo que, en base a la información disponible, predijera los resultados de la Encuesta PACES en los establecimientos del país. Por último, estos resultados se compararon con datos reales de un grupo de casos que no le habían sido suministrados al algoritmo.

Se utilizaron **herramientas** de programación en Python, trabajando los modelos en Jupyter Notebook, que permite crear y compartir documentos que contienen código, ecuaciones, visualizaciones y texto narrativo. Sus usos incluyen limpieza y transformación de datos,

simulaciones, modelamiento estadístico, visualización, aprendizaje automático y más ⁷. Las bibliotecas utilizadas fueron:

- **Pandas:** es una biblioteca de software escrita como extensión de NumPy para manipulación y análisis de datos para el lenguaje de programación Python. Ofrece estructuras de datos y operaciones para manipular tablas numéricas y series temporales. Es un software libre.
- **NumPy:** es una biblioteca de Python que proporciona un objeto de matriz multidimensional, varios objetos derivados (como matrices y matrices enmascaradas) y una variedad de rutinas para operaciones rápidas en matrices⁸.
- **Scikit-learn:** es una biblioteca que implementa distintos algoritmos para construir modelos y de la cual se utilizaron:
 - **Regresión logística:** la regresión logística es una variante de la regresión lineal que se utiliza en el contexto de clasificación, en especial en procesos cuyo resultado adopta dos valores excluyentes (“cara” o “sello” por ejemplo).
 - **Árboles de decisión:** son representaciones gráficas de posibles soluciones a una decisión basadas en ciertas condiciones. Los árboles de decisión tienen un primer nodo llamado raíz y luego se descomponen en el resto de los atributos de entrada en dos ramas planteando una condición que puede ser cierta o falsa. Se bifurca cada nodo en 2 y vuelven a subdividirse hasta llegar a las hojas que son los nodos finales y que equivalen a respuestas a la solución (Bagnato, 2020).
 - **Random forest:** es un tipo de Ensamble en Machine Learning en donde se combinan diversos árboles y la salida de cada uno se contará como “un voto” y la opción más votada será la respuesta del Bosque Aleatorio (Bagnato, 2020). Las características de los árboles utilizados para el bosque son diversas.
- **XGboost:** es una biblioteca optimizada de ensamble de árboles con aumento de gradiente⁹. XGBoost proporciona un refuerzo de árbol paralelo (también conocido como GBDT, GBM) que resuelve muchos problemas de ciencia de datos de una manera rápida y precisa¹⁰.
- **Catboost:** es un algoritmo de aprendizaje automático que utiliza un aumento de gradiente basado en árboles de decisión y está disponible como una biblioteca de código abierto.

⁷ Página web Project Jupyter: <https://jupyter.org/>

⁸ Página web NumPy: <https://numpy.org/>

⁹ Modelos formados por un conjunto de árboles de decisión individuales entrenados de forma secuencial, de forma que cada nuevo árbol intenta mejorar los errores de los árboles anteriores. La predicción de una nueva observación se obtiene agregando las predicciones de todos los árboles individuales que forman el modelo (Amat, 2020).

¹⁰ XGBoost disponible en: <https://xgboost.readthedocs.io/en/latest/>

Catboost permite usar variables no numéricas en lugar de preprocesar datos dedicando tiempo y esfuerzo en convertirlos en números. El algoritmo es usado en sistemas de recomendación, asistentes personales, autos autónomos, predicción de clima, entre otros.¹¹

3.1 Exploración de datos

En esta etapa se identificó las fuentes de datos administrativos disponibles para abordar los objetivos del estudio. La investigación usó datos del Programa PACES-PUCV, datos administrativos de las instituciones que componen el Sistema de Aseguramiento de la Calidad de la Educación y datos de otros organismos públicos. La Tabla N°5 expone las bases de datos utilizadas:

Tabla N°5. Fuentes de datos utilizadas.

Base de datos	Fuente	Año
Base estudiantes 2016 ¹²	Programa de Apoyo a la Convivencia Escolar (PACES)	2016
Base estudiantes 2018	Programa de Apoyo a la Convivencia Escolar (PACES)	2018
Cuestionarios de Calidad y Contexto de la Educación	Agencia de Calidad de la Educación	2016, 2018
Base Indicadores de Desarrollo Personal y Social	Agencia de Calidad de la Educación	2016, 2018
Base Categorías de Desempeño educación básica	Agencia de Calidad de la Educación	2016, 2018
Base de denuncias por RBD, Región y comuna	Superintendencia de Educación	2016, 2018
Base de denuncias sancionadas por RBD, Región y comuna.	Superintendencia de Educación	2016, 2018
Prioridades 2016 con IVE SINAE básica, Media y Comunal.	Junta Nacional de Auxilio Escolar y Becas	2016, 2018
Resumen Prioritarios y Beneficiarios SEP por Establecimiento Educacional	Ministerio de Educación	2016, 2018

Fuente: Elaboración propia.

En cuanto a Cuestionarios de Calidad y Contexto de la Educación, se trabajó con las “Glosas” SIMCE 2016 y 2018 para consolidar una base de datos con variables equivalentes, ya que muchas de ellas cambiaron su código de identificación entre ambos procesos.

Las variables 2018 “sin equivalencia” en 2016 no fueron incluidas en la base de datos del estudio (aunque apuntaban directamente a las características de la convivencia en la escuela). El término “sin equivalencia” indica solamente que la variable no estuvo presente en los Cuestionarios de

¹¹ Catboost disponible en: <https://catboost.ai/>

¹² Las bases de datos de PACES 2016 y 2018 contienen las respuestas de estudiantes de quinto, sexto, séptimo y octavo básico. Se mantuvieron todos los cursos con el propósito de no disminuir la muestra de entrenamiento de los modelos.

Calidad y Contexto del año 2016. El siguiente cuadro muestra los nombres de las variables 2016 utilizadas junto a sus equivalentes en 2018:

Tabla N° 6. Variables SIMCE 2018-2016 y sus equivalencias

Variable SIMCE 2018	Variable SIMCE 2016	Pregunta
cest_p21_01	cest_p37_01	Me gusta mucho mi escuela
cest_p21_02	cest_p37_02	Me siento orgulloso de mi escuela
cest_p09_01	cest_p11_03	Los estudiantes de mi curso nos tratamos con respeto
cest_p22_01	cest_p12_03	Mis profesores me hacen sentir que soy parte importante de mi escuela
cest_p22_02	cest_p12_04	El director me hace sentir que soy parte importante de mi escuela
cest_p37_02	cest_p10_02	Corrige las tareas que envié para la casa
cest_p37_03	cest_p10_03	Vuelve a explicar una materia si algún estudiante se lo pide
cest_p05_01	cest_p09_01	Mis profesores me dicen que soy buen estudiante
cest_p05_02	cest_p09_02	Mis profesores dicen que soy capaz de aprender
cest_p11_01	Sin equivalencia	Las normas de convivencia son conocidas por todos los estudiantes
cest_p11_02	Sin equivalencia	Las normas de la escuela son claras
cest_p11_03	Sin equivalencia	Los estudiantes respetamos las normas de convivencia
cest_p11_04	Sin equivalencia	Los profesores nos enseñan las consecuencias de no cumplir las normas
cest_p18_01	cest_p19_01	Los profesores de mi curso piden nuestra opinión en clases
cest_p18_02	cest_p19_02	En mi curso conversamos los problemas para buscar soluciones
cest_p18_03	Sin equivalencia	Mis compañeros resuelven sus problemas conversando
cest_p18_04	cest_p19_03	Mi opinión es tomada en cuenta por mis profesores
cest_p18_05	cest_p19_04	Mis profesores me animan a decir mis opiniones
cest_p18_06	cest_p19_05	Puedo expresar mi opinión, aunque sea distinta a la de mis compañeros
cest_p10_01	cest_p11_02	Los estudiantes de mi curso mantenemos limpia la sala de clases
cest_p10_02	Sin equivalencia	Mis compañeros cuidan que la escuela esté limpia (por ejemplo, baños, patio, canchas, etc.)
cest_p12_03	cest_p13_03	Mis profesores explican por qué es importante cumplir las normas
cest_p03_01	cest_p04_02	Puedo hacer bien las tareas y trabajos difíciles
cest_p03_02	cest_p04_03	Puedo aprender con facilidad
cest_p03_03	cest_p04_04	Sé que puedo tener buenas notas
cest_p03_04	cest_p04_05	Sé que soy buen estudiante
cest_p03_05	cest_p04_06	Soy capaz de termina el año con un buen promedio de notas

Fuente: Elaboración propia.

Dentro del catastro de recursos ausentes, es importante destacar que las subdimensiones de la Encuesta PACES contienen factores escolares que no están presentes en los Cuestionarios de Calidad y Contexto de la Educación. Por ejemplo, la subdimensión de “apoyo” consulta de forma explícita si los profesores favorecen el buen trato y la convivencia escolar, el conocimiento de las reglas de convivencia o la preocupación de los profesores por los estudiantes. De forma similar, la dimensión “victimización” considera como subdimensiones la victimización verbal, ciber, física, por exclusión social y por acoso sexual, con un total de 28 variables.

3.2 Preparación de datos

3.2.1 Cálculo de resultados PACES a nivel de establecimiento

Se calculó los resultados de cada variable, a nivel de escuela, siguiendo las reglas de la Encuesta PACES a estudiantes. Por ejemplo, para obtener el resultado a nivel de escuela: 1) se dividió cada respuesta individual por tres y 2) para el cálculo de la dimensión se promediaron los ítems correspondientes. Esta tarea se automatizó en Excel según la regla aplicable en cada caso.

3.2.2 Construcción de base de datos

A partir de las anteriores fuentes de datos se construyó la base de datos de este estudio, usando como “llave” los RBD de los establecimientos educacionales incluidos en la “Base de datos 2016” y la “Base de datos 2018” de la Encuesta del Programa PACES-PUCV. La base final conservó 1346 RBD de un total de 1441 escuelas presentes en las bases 2016 y 2018 (equivalente a 97,57% de los registros). El criterio para excluir algunas escuelas fue la baja cantidad de estudiantes que respondieron la encuesta, lo que podría sesgar los resultados calculados a nivel de establecimiento¹³. A continuación, se explica brevemente el procedimiento para construir la base de datos para entrenamiento y prueba de los algoritmos de regresión y clasificación:

- Se calculó el resultado a nivel de escuela a partir de las respuestas de los estudiantes en cada dimensión y subdimensión de la Encuesta PACES. El procedimiento que se siguió en cada caso dependió de la escala de respuesta utilizado por la pregunta. A modo de ejemplo, para obtener el resultado a nivel de escuela para las variables relacionadas a la subdimensión “satisfacción”, se debe dividir la respuesta individual de cada variable por tres, promediar los resultados correspondientes y multiplicar por cien. Luego, para obtener el resultado a nivel de subdimensión se promedian los resultados de las variables correspondientes.
- Para obtener el resultado a nivel de dimensión, por ejemplo, “clima escolar”, se promedian los resultados de las subdimensiones, obteniendo así un puntaje de “clima escolar” para el establecimiento. Este cálculo y el anterior se automatizaron en Excel.
- Luego, se calculó el resultado a nivel de establecimiento educacional a partir de las respuestas de los estudiantes en los Cuestionarios de Calidad y Contexto de la Educación 2016 y 2018. Al igual que en el caso de PACES, el procedimiento utilizado dependió de la escala de respuesta que utilizaba la pregunta.

¹³ Si un número muy pequeño de estudiantes responden la encuesta en una escuela, sus respuestas equivaldrían al resultado del establecimiento con lo cual estos estudiantes quedarían sobre representados introduciendo una distorsión indeseable en los resultados.

- Calculados los resultados a nivel de escuela se generó la base de datos final usando como “llave” el RBD del establecimiento.
- En cuanto a las variables categóricas, fueron incluidas el grupo socioeconómico del establecimiento, ruralidad, la dependencia de la escuela, su categoría de desempeño y la etiqueta según su tipo de convivencia.

3.2.3 Análisis exploratorio de datos

A continuación, se resumen las acciones y resultados del análisis exploratorio de datos:

- Se identifican casos de datos perdidos en algunas variables. En este punto, destaca el caso del IDPS ICCE donde se contabilizaron 40 establecimientos sin resultado en la medición, según la Agencia de Calidad de la Educación esto se debe a que, para asegurar confiabilidad, el cálculo solo se realiza cuando cierto porcentaje de los cuestionarios fueron respondidos.
- En los casos anteriores, para evitar disminuir más la muestra de establecimientos, se imputó el valor perdido con la media de la variable correspondiente. En el caso de variables categóricas con datos perdidos se utilizó la moda de la variable correspondiente.
- La recodificación de las variables categóricas para alimentar los algoritmos se realizó en Jupyter Notebook, cuando así se requirió.
- Los valores atípicos se identificaron a través de la inspección visual de gráficos de dispersión y de cajas. Estos normalmente correspondían a establecimientos con un nivel bajo de estudiantes participantes, lo que, como se señaló previamente introducía un sesgo en los resultados del establecimiento, por lo cual no fueron considerados en los modelos.
- Al analizar la base de datos final se encontraron correlaciones altas y directas entre las variables “clima de aula” y “clima escolar” de la Encuesta PACES junto a correlaciones moderadas e inversas entre dichas variables y las dimensiones “victimización entre pares” y “seguridad en la escuela”. Además, se evidencia una correlación moderada entre “clima de aula” y “clima escolar” y los IDPS “clima de convivencia escolar” y “participación y formación ciudadana”. También se constatan correlaciones moderadas y directas con variables de las Cuestionarios de Calidad y Contexto como `cest_p21_01` (“Me gusta mucho mi escuela”), `cest_p21_02` (“Me siento orgulloso de mi escuela”), entre otras.
- Los Cuestionarios de Calidad y Contexto de la Educación contienen una gran cantidad de variables. Por ejemplo, el cuestionario a estudiantes, 6° básico, 2018, contenía 158 variables. Por ello, se filtraron variables a utilizar seleccionando las pertinentes en función de la literatura revisada. Así, las variables utilizadas Cuestionarios de Calidad y Contexto fueron las siguientes para los años 2016 y 2018 respectivamente:

Tabla N°7. Variables utilizadas Cuestionarios de Calidad y Contexto SIMCE 2016

Dimensión PACES	Subdimensión PACES	Factores según marco teórico	Variable SIMCE	Pregunta SIMCE
Clima escolar	Satisfacción	Factores individuales	cest_p37_01 cest_p37_02	<ul style="list-style-type: none"> • Me gusta mucho mi escuela. • Me siento orgulloso de mi escuela.
	Apoyo	Factores escolares	cest_p09_01 cest_p09_02 cest_p12_03	<ul style="list-style-type: none"> • Mis profesores me dicen que soy buen estudiante. • Mis profesores me dicen que soy capaz de aprender. • Mis profesores me hacen sentir que soy parte importante de mi escuela.
	Participación	Factores escolares	cest_p19_01 cest_p19_03 cest_p19_04 cest_p19_05 cest_p19_06	<ul style="list-style-type: none"> • Los profesores de mi curso piden nuestra opinión en clases. • Mi opinión es tomada en cuenta por mis profesores. • Mis profesores me animan a decir mis opiniones. • Sé que puedo expresar mi opinión, aunque sea distinta a la de mis compañeros. • Los estudiantes participamos en la organización de las actividades del curso (por ejemplo, actos, presentaciones, paseos, etc.)
	Percepción del ambiente físico	Factores escolares	cest_p11_02 cest_p12_02	<ul style="list-style-type: none"> • Los estudiantes de mi curso mantenemos limpia la sala. • Los estudiantes cuidan que la escuela esté limpia (por ejemplo, los baños, patios, cancha, etc.).
	Organización y Participación del curso	Factores escolares	cest_p13_03	<ul style="list-style-type: none"> • Mis profesores explican por qué es importante cumplir las normas.
Clima de aula	Relación entre pares	Factores escolares	cest_p11_03 cest_p19_02	<ul style="list-style-type: none"> • Los estudiantes de mi curso nos tratamos con respeto. • En mi curso conversamos los problemas para buscar soluciones.
	Orientación al aprendizaje y altas expectativas	Factores individuales	cest_p04_01 cest_p04_02 cest_p04_03 cest_p04_04 cest_p04_05 cest_p04_06	<ul style="list-style-type: none"> • Puedo hacer las tareas, aunque sean difíciles. • Sé que siempre puedo hacer bien mis tareas. • Puedo aprender con facilidad y rapidez. • Me puedo sacar buenas notas si me esfuerzo. • Sé que soy un buen estudiante. • Soy capaz de terminar el año con un buen promedio de notas.
Victimización	Victimización entre pares	Factores escolares	cest_p16_01 cest_p16_02	<ul style="list-style-type: none"> • Algún compañero se ha burlado de ti. • Algún compañero te ha pegado.

Fuente: Elaboración propia.

Como se observa en la Tabla N°7, las variables utilizadas fueron agrupadas en función de su relación con las subdimensiones de la Encuesta PACES y los factores que incluyen en la convivencia escolar según la revisión de literatura (punto 2.1 del documento). La misma situación ocurre en el caso de la Tabla N°8 que se presenta a continuación:

Tabla N°8. Variables utilizadas Cuestionarios de Calidad y Contexto SIMCE 2018

Dimensión PACES	Subdimensión PACES	Factores según marco teórico	Variable SIMCE	Pregunta SIMCE
Clima escolar	Satisfacción	Factores individuales	cest_p21_01 cest_p21_02	<ul style="list-style-type: none"> • Me gusta mucho mi escuela. • Me siento orgulloso de mi escuela.
	Apoyo	Factores escolares	cest_p05_01 cest_p05_02 cest_p22_01 cest_p22_02	<ul style="list-style-type: none"> • Mis profesores me dicen que soy buen estudiante. • Mis profesores me dicen que soy capaz de aprender. • Mis profesores me hacen sentir que soy parte importante de mi escuela. • El profesor revisa que todos hayamos hecho las tareas.
	Participación	Factores escolares	cest_p18_01 cest_p18_04 cest_p18_05 cest_p18_06	<ul style="list-style-type: none"> • Los profesores de mi curso piden nuestra opinión en clases. • Mi opinión es tomada en cuenta por mis profesores. • Mis profesores me animan a decir mis opiniones. • Sé que puedo expresar mi opinión, aunque sea distinta a la de mis compañeros.
	Percepción del ambiente físico	Factores escolares	cest_p10_01 cest_p10_02	<ul style="list-style-type: none"> • Los estudiantes de mi curso mantenemos limpia la sala • Los estudiantes cuidan que la escuela esté limpia (por ejemplo, los baños, patios, cancha, etc.).
Clima de aula	Organización y Participación del curso	Factores escolares	cest_p12_03	<ul style="list-style-type: none"> • Mis profesores explican por qué es importante cumplir las normas.
	Relación entre pares	Factores escolares	cest_p09_01 cest_p18_02	<ul style="list-style-type: none"> • Los estudiantes de mi curso nos tratamos con respeto • En mi curso conversamos los problemas para buscar soluciones.
	Orientación al aprendizaje y altas expectativas	Factores individuales	cest_p03_01 cest_p03_02 cest_p03_03 cest_p03_04 cest_p03_05	<ul style="list-style-type: none"> • Puedo hacer bien las tareas y trabajos difíciles. • Puedo aprender con facilidad. • Sé que puedo sacarme buenas notas. • Sé que soy un buen estudiante. • Soy capaz de terminar el año con un buen promedio de notas.
Victimización	Victimización entre pares	Factores escolares	cest_p16_01 cest_p16_02	<ul style="list-style-type: none"> • Algún compañero se ha burlado de ti. • Algún compañero te ha pegado.

Fuente: Elaboración propia.

Como se observa, las variables disponibles corresponden a “factores escolares” y “factores individuales” de acuerdo a la literatura revisada. En cuanto a los “factores familiares” “contextuales” y de “política pública” que pueden impactar en la convivencia y comportamientos, dicha información no está disponible.

En el caso de las denuncias a la Superintendencia de Educación, se consideraron exclusivamente aquellas relacionadas con la convivencia escolar, incluyéndose en el listado principalmente las siguientes categorías: a) maltrato físico de adulto a alumno (a); b) maltrato físico entre alumnos (as); c) maltrato psicológico entre alumnos(as); d) maltrato de apoderado (a) a docente y/o asistente de la educación; e) maltrato psicológico de adulto a alumno (a) y f) maltrato de alumno (a) a docente y/o asistente de la educación. De acuerdo a la revisión de literatura estas variables se encuentran dentro de los “factores escolares” que influyen la convivencia escolar.

3.3 Planificación del modelo

En esta etapa se determinó qué modelos se utilizarían para tareas de regresión y qué modelos para clasificación. En el primer caso se escogió la regresión múltiple, los árboles de regresión, bosques de regresión, *K-nearest neighbors* y *XGboost*. En el segundo caso se escogió la Regresión Logística, árboles de clasificación, bosques de clasificación, *Catboost* y *XGboost*.

3.4 Elaboración del modelo

Esta etapa incluye las siguientes actividades: a) utilizar los conjuntos de datos para pruebas, entrenamiento y producción; b) evaluar la viabilidad y confiabilidad de los datos para usar en los modelos predictivos. y c) Finalmente, desarrollar, probar y perfeccionar los modelos analíticos (Schmarzo, 2013). Para Schmarzo (2013) esta es una etapa altamente iterativa, en que se fallará varias veces al probar diferentes variables y técnicas de modelado. Se ejecutará en las secciones 4.1 y 4.2 del documento.

3.5 Comunicar resultados

En esta etapa se incluyen las siguientes actividades: a) Determinar la calidad y confiabilidad del modelo analítico y la significancia estadística, mensurabilidad y capacidad de acción de los conocimientos analíticos resultantes y b) Desarrollar los cuadros y gráficos para comunicar los conocimientos, los resultados y las recomendaciones del modelo analítico (Schmarzo, 2013). Se ejecutará en la sección 4.3

3.6 Uso

Se ejecutará en la sección 4.5.

4 Resultados

Se exponen los resultados de los modelos de regresión y clasificación para predecir el valor de la Encuesta PACES en las escuelas del país. De las escuelas con Encuesta PACES ($n=1.346$), se utilizó un total de 1.010 como muestra de entrenamiento (equivalente a un 75% del total de escuelas) y 336 como muestra de validación (equivalente a un 25% del total de escuelas).

A través de los modelos de regresión se buscó predecir el resultado numérico y continuo de las dimensiones “clima de aula” y “clima escolar”, obteniendo resultados moderados tanto para la muestra de entrenamiento como para la muestra de validación. Como expone la Tabla N°9, los “bosques de regresión” obtuvieron el mejor resultado, que alcanzó un R^2 de 69,4% en entrenamiento y 48,6% en validación para la variable “clima de aula” y un R^2 de 67,4% en entrenamiento y 45,9% en validación para la variable “clima escolar”. Por su parte, la Tabla N°10 muestra los coeficientes de determinación ajustados (R^2 ajustados) para el set de entrenamiento. Estos resultados no son del todo fuertes lo que se analiza en el punto 4.1.

Con los modelos de clasificación se buscó predecir un resultado discreto, que fue definido como una “buena” o “débil” convivencia escolar en las escuelas. Se revisó el resultado de los algoritmos usando diferentes puntos de corte para ambas clases, concretamente se utilizó la media y las siguientes desviaciones estándar respecto ella de las variables “clima escolar” y “clima de aula”: $\sigma=0,2$, $\sigma=0,3$ y $\sigma=0,4$. El modelo de Regresión Logística obtuvo un *accuracy* del 78,51%, logrando una *precisión* del 66,1% y un *recall* del 81,25% para la clase “débil” convivencia¹⁴. Esto significa que el modelo detecta correctamente dicha clase y, aunque existe margen de mejora, es capaz de disminuir los “falsos negativos” en mayor proporción que el resto de los modelos entrenados. Estos resultados se analizan en el punto 4.2.

4.1 Modelos de Regresión

La predicción de los resultados de la Encuesta PACES con modelos de regresión se utiliza para variables dependientes cuantitativas continuas, concretamente para los resultados de las variables (o subdimensiones) “clima de aula” y “clima escolar”. Para ello se utilizaron modelos de regresión múltiple, árboles de regresión, bosques de regresión, *k-nearest neighbors* y *XGboost*.

Los resultados se presentan en la siguiente tabla:

¹⁴ Dado el desbalance natural del fenómeno, se usan como métricas de desempeño de los modelos el *recall* y *precision* de ambas clases. El *recall* es un valor entre 0 y 1, que es la tasa de acierto del clasificador binario de una clase determinada respecto al total de casos que efectivamente pertenecen a dicha clase. La tasa de falsos negativos es $1 - \text{recall}$. Esta tasa es importante, porque al minimizar los falsos negativos, permite asegurar que todos los estudiantes que potencialmente pudieran abandonar reciban acompañamiento y no queden indetectados. Por su parte la *precision* es un valor entre 0 y 1, que es la tasa de acierto respecto al total de predicciones que el clasificador binario hace para una clase determinada. La tasa de falsos positivos es $1 - \text{precision}$. Por último, F1 es media armónica entre el *precision* y el *recall* de una clase.

Tabla N°9. Resultado modelos de regresión “clima de aula”, “clima escolar”

Modelo		R² “Clima de Aula”	R² “Clima Escolar”
<i>Regresión Múltiple</i>	Datos de entrenamiento	42,03%	44,48%
	Datos de prueba	41,70%	44,31%
<i>Árbol de regresión</i>	Datos de entrenamiento	45,6%	47%
	Datos de prueba	28,2%	34,6%
<i>Bosques de regresión</i>	Datos de entrenamiento	69,4%	67,4%
	Datos de prueba	48,6%	45,9%
<i>K-nearest neighbors</i>	Datos de entrenamiento	54,8%	51,8%
	Datos de prueba	42,5%	41,7%
<i>XGboost</i>	Datos de entrenamiento	73,43%	70,43%
	Datos de prueba	46,06%	39,70%

Fuente: Elaboración propia

Por su parte, los coeficientes de determinación ajustados para el set de entrenamiento se muestran a continuación:

Tabla N°10. Resultado modelos de regresión “clima de aula”, “clima escolar” (R² ajustados)

Modelo	R² “Clima de Aula”	R² “Clima Escolar”
<i>Regresión Múltiple</i>	41,68%	44,20%
<i>Árbol de regresión</i>	40,79%	28,9%
<i>Bosques de regresión</i>	55,3%	57,6%
<i>K-nearest neighbors</i>	50,45%	45,75%
<i>XGboost</i>	60,76%	55,78%

Fuente: Elaboración propia

La imposibilidad de conseguir la predicción de la variable continua podría deberse a que las subdimensiones de PACES miden ámbitos de la convivencia escolar que no estuvieron representados en el modelo, pues su medición no estaba disponible como datos administrativos, situación que se analizará en el apartado de limitaciones de la investigación.

4.2 Modelos de Clasificación

Para clasificar las escuelas en función de su convivencia escolar se utilizó los datos de entrada de las escuelas (n=1.346) en cinco modelos capaces de arrojar resultados discretos. En este caso la clasificación fue binaria catalogando las escuelas en los grupos de “buena” o “débil” convivencia escolar. Como se observará más adelante, se revisó el desempeño de los algoritmos usando diferentes puntos de corte para las clases “buena” y “débil” convivencia escolar (se utilizó la media y las siguientes desviaciones estándar respecto a la media de las variables “clima escolar” y “clima de aula” de la Encuesta PACES: $\sigma=0,2$, $\sigma=0,3$, $\sigma=0,4$).

En todos modelos se dividió los datos de entrada en una muestra de entrenamiento y una muestra de validación o prueba. Ésta última con la finalidad de testear el modelo con datos desconocidos

por el algoritmo. Del conjunto de datos disponibles, usamos un 75% aleatoriamente para entrenamiento y un 25% como datos de prueba. A través de la comparación de la categoría real de los datos de validación y la predicha obtenemos los aciertos del modelo.

En cada uno de los puntos de corte propuestos, las escuelas etiquetadas con “débil” convivencia escolar conformaron la clase minoritaria. Este desbalance en las muestras puede perjudicar el desempeño del modelo al mostrar buenos resultados acertando solo a la clase mayoritaria, en lugar de identificar adecuadamente la clase de “débil” convivencia. A fin de contrarrestar el desbalance de las muestras usadas para entrenar el algoritmo, se ajustaron los parámetros de los modelos penalizando a la clase mayoritaria.

Adicionalmente, se utilizó validación cruzada¹⁵ en todos los modelos (k-folds=5) para evitar el sobreajuste a los datos de entrenamiento. Si bien este procedimiento normalmente produce un ajuste a la baja en las métricas de resultado, con esto se resguarda que el desempeño de los modelos (sea este positivo o negativo) no sea producto de la selección aleatoria de las muestras por parte del algoritmo. Se tomó esta opción considerando que la muestra de establecimientos educacionales es de tamaño medio (n=1.346).

4.2.1 Resumen y evaluación de modelos de clasificación

Como se ha indicado, al tratarse de un problema de aprendizaje supervisado se debe contar con las etiquetas de los establecimientos educacionales según su convivencia escolar.

Para etiquetar los establecimientos con “débil” convivencia escolar, se utilizó como primer punto de corte la media para ambas variables de la Encuesta PACES (“clima escolar” y “clima de aula”). Las escuelas que se encuentran bajo ese límite se consideran con “débil” convivencia escolar, obteniendo 607 establecimientos educacionales en la clase “débil” y 739 en la clase “buena”. La Tabla N°11 muestra los resultados de clasificación obtenidos con este punto de corte usando validación cruzada (k-folds=5):

Tabla N°11. Resultados por clase con validación cruzada (\bar{X})

Modelo		Accuracy	Precision	Recall	F1-score	Support
Regresión Logística	Buena	76,54%	81,82%	77,27%	79,48%	198
	Débil		70%	75,54%	72,66%	139
Árbol de Clasificación	Buena	63,5%	69,39%	68,69%	69,04%	198
	Débil		56,03%	56,83%	56,43%	139
Random Forest	Buena	74,48%	74,44%	83,84%	78,86%	198

¹⁵ K-folds es una herramienta de validación cruzada que consiste en dividir los datos de la muestra en “k” subconjuntos, usando uno de los subconjuntos como datos de prueba y el resto como datos de entrenamiento tantas iteraciones como se haya definido. En este caso se definió un número de cinco iteraciones (k-folds=5) y los resultados entregados representan la media aritmética de las cinco iteraciones.

	Débil		71,93%	58,99%	64,82%	139
<i>Catboost</i>	Buena	75,07%	76,9%	82,3%	79,5%	198
	Débil		72%	64,7%	78,2%	139
<i>XGboost</i>	Buena	71,09%	75%	77,27%	76,12%	198
	Débil		66,17%	63,31%	64,71%	139

Fuente: Elaboración propia.

El mejor resultado con este punto de corte corresponde a la Regresión Logística. El *accuracy* alcanzado es de un 76,54%, es decir, la clasificación del modelo es correcta en un 76,54% de las predicciones. Al observar su *precision* por clase, el resultado es de un 70% para los establecimientos con “débil” convivencia escolar y de un 81,82% para los establecimientos con “buena” convivencia escolar. Además, el *recall* nos muestra que el modelo reconoce equilibradamente ambas clases, con un 77,27% para la clase con “buena” convivencia escolar y un 75,54% para la clase con “débil” convivencia escolar.

La Tabla N°12 muestra los resultados de los modelos usando como punto de corte de clasificación la media de las variables indicadas menos 0,2 desviaciones estándar. Con este punto de corte se obtienen 488 establecimientos educacionales en la clase “débil” y 858 en la clase “buena”:

Tabla N°12. Resultados por clase con validación cruzada (\bar{x} menos $\sigma=0,2$)

Modelo		<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	F1-score	Support
<i>Regresión Logística</i>	Buena	78,33%	88,6%	77,03%	82,41%	222
	Débil		64,58%	80,87%	71,81%	115
<i>Árbol de Clasificación</i>	Buena	67,06%	77,13%	77,48%	77,3%	222
	Débil		56,14%	55,65%	55,9%	115
<i>Random Forest</i>	Buena	79,22%	80%	90,09%	84,75%	222
	Débil		74,71%	56,52%	64,36%	115
<i>Catboost</i>	Buena	77,74%	82,4%	84,2%	83,3%	222
	Débil		68,2%	65,2%	66,7%	115
<i>XGboost</i>	Buena	72,13%	82,35%	81,98%	82,17%	222
	Débil		65,52%	66,09%	65,8%	115

Fuente: Elaboración propia.

El mejor resultado con este punto de corte corresponde a la Regresión Logística, que es también el mejor resultado obtenido en general. El *accuracy* con alcanzado es de un 78,33%, es decir, la clasificación del modelo es correcta en un 78,33% de las predicciones. Al observar su *precision* por clase, el resultado es de un 64,58% para los establecimientos con “débil” convivencia escolar y de un 88,6% para los establecimientos con “buena” convivencia escolar. Además, el *recall* nos muestra que el modelo reconoce equilibradamente ambas clases, con un 77,03% para la clase con “buena” convivencia escolar y un 80,86% para la clase con “débil” convivencia escolar.

Si bien la *precision* del modelo disminuye respecto al punto de corte anterior, obtenemos un *recall* cercano al 81%, que nos permite identificar los establecimientos con “débil” convivencia. Las métricas presentadas para este modelo se consiguen de la siguiente Matriz de Confusión:

Figura N°6. Matriz de Confusión

Etiqueta verdadera	0	171	51
	1	22	93
		0	1
		Etiqueta predicha	

Fuente: Elaboración propia.

El *recall* de la clase “débil” convivencia es igual al total de verdaderos positivos (es decir, escuelas con “débil” convivencia, clasificados con “débil” convivencia por el modelo) dividido en la suma de los verdaderos positivos más los falsos positivos de la clase “buena” convivencia (es decir, establecimientos que teniendo “débil” convivencia fueron clasificados con “buena” convivencia). En este caso, el cálculo es $93/(93+22)=80,87\%$ y corresponde al mejor *recall* de todos los modelos entrenados, que nos permite identificar la clase “débil”.

El error inverso, es decir, clasificar escuelas con “buena” convivencia como si tuvieran una convivencia “débil” es aceptable, mientras se busca reducir al mínimo las escuelas con “débil” convivencia clasificados con la etiqueta “buena” convivencia. Es por eso que una métrica clave para la clase de convivencia “débil” es el *recall*, ya que nos interesa disminuir al mínimo los “falsos negativos”, es decir, escuelas que teniendo “débil” convivencia, son clasificados con “buena” convivencia. El modelo obtuvo un *recall* superior al 80%.

La Tabla N°13 muestra los resultados usando como punto de corte de clasificación la media de las variables indicadas menos 0,3 desviaciones estándar. Usando este punto de corte se obtienen 423 establecimientos educacionales en la clase “débil” y 923 en la clase “buena”.

Tabla N° 13. Resultados por clase con validación cruzada (\bar{x} menos $\sigma=0,3$)

Modelo		Accuracy	Precision	Recall	F1-score	Support
<i>Regresión Logística</i>	Buena	75,36%	86,8%	75%	80,47%	228
	Débil		59,29%	76,15%	66,67%	109
<i>Árbol de Clasificación</i>	Buena	67,65%	77,93%	75,88%	76,89%	228
	Débil		52,17%	55,05%	53,57%	109
<i>Random Forest</i>	Buena	75,09%	80,08%	89,91%	84,71%	228
	Débil		71,6%	53,21%	61,05%	109
<i>Catboost</i>	Buena	77,75%	84,3%	82,5%	83,4%	228
	Débil		64,9%	67,9%	66,4%	109
<i>XGboost</i>	Buena	73,69%	82,25%	83,33%	82,79%	228
	Débil		64,15%	62,39%	63,26%	109

Fuente: Elaboración propia.

La Tabla N°14 muestra los resultados usando como punto de corte de clasificación la media de las variables indicadas menos 0,4 desviaciones estándar. Usando este punto de corte se obtienen 375 establecimientos educativos en la clase “débil” y 971 en la clase “buena”.

Tabla N°14. Resultados por clase con validación cruzada (\bar{x} menos $\sigma=0,4$)

Modelo		Accuracy	Precision	Recall	F1-score	Support
<i>Regresión Logística</i>	Buena	75,65%	89%	74,79%	81,28%	238
	Débil		56,2%	77,78%	65,25%	99
<i>Árbol de Clasificación</i>	Buena	69,13%	80,25%	80,25%	80,25%	238
	Débil		52,53%	52,53%	52,53%	99
<i>Random Forest</i>	Buena	76,85%	82,38%	90,34%	86,17%	238
	Débil		69,74%	53,54%	60,57%	99
<i>Catboost</i>	Buena	77,45%	85,2%	82,4%	83,8%	238
	Débil		60,7%	65,7%	63,1%	99
<i>XGboost</i>	Buena	73,85%	80,95%	85,71%	83,27%	238
	Débil		60%	51,52%	55,43%	99

Fuente: Elaboración propia.

4.3 Importancia de las variables para el modelo de clasificación

Escogida la Regresión Logística y el punto de corte señalado, en este apartado se muestra la contribución relativa de cada variable al resultado del modelo de clasificación. A través de los coeficientes de la regresión se obtiene la influencia de las variables en la clasificación. En ese sentido, la importancia de las variables se muestra en la Tabla N° 15:

Tabla N° 15. Importancia de las variables para el modelo Regresión Logística.

	feature	feature_importance
8	cest_p16_02	1.351288
17	cest_p12_03	1.240704
11	cest_p05_01	0.901170
5	cest_p21_02	0.789534
23	cat_des_INSUFICIENTE	0.610795
6	cest_p09_01	0.599057
15	cest_p18_05	0.566930
10	cest_p22_02	0.544247
12	cest_p18_01	0.478290
25	cat_des_MEDIO-BAJO	0.383194
14	cest_p18_04	0.358882
21	cod_rural_Urbano	0.358826
19	cod_grupo_Medio alto	0.297231
13	cest_p18_02	0.285647
16	cest_p18_06	0.279294
1	ind_cc	0.203354
18	cod_grupo_Medio	0.190088
20	cod_grupo_Medio bajo	0.169137
22	cod_depe_Part_sub	0.149551
7	cest_p16_01	0.111861
4	cest_p21_01	0.109206
9	cest_p22_01	0.087838
0	ind_am	0.087486
2	ind_hv	0.060711
3	ind_pf	0.056361
24	cat_des_MEDIO	0.027948

Fuente: Elaboración propia.

Como se observa en la Tabla N°15 las variables de mayor influencia en el modelo son **cest_p16_02** (“algún compañero te ha pegado”), **cest_p12_03** (“mis profesores explican por qué es importante cumplir las normas”), **cest_p05_01** (“mis profesores me dicen que soy buen estudiante”), **cest_p21_02** (“me siento orgulloso de mi escuela”), la categoría de desempeño “insuficiente”, **cest_p09_01** (“los estudiantes de mi curso nos tratamos con respeto”), **cest_p18_05** (“mis profesores me animan a decir mis opiniones”) y **cest_p22_02** (“el director me hace sentir que soy parte importante de mi escuela”).

4.4 Clasificación nacional usando el modelo de Regresión Logística

Con el propósito de clasificar las escuelas del país según su “débil” o “buena” convivencia escolar, se construyó una nueva base de datos. Esta base abarcó todos los establecimientos de educación básica del país en el año 2018. Del total de escuelas (n=7.322) se descartaron 483 particulares pagadas y diecinueve de grupo socioeconómico alto, ya que los datos de la Encuesta PACES no incluían establecimientos con esas características, por lo que los modelos no se habían entrenado con casos similares. De esta forma, se generó una base con 6.820 escuelas de educación básica (equivalente a un 93,14% del total nacional), que, en base a sus atributos incluidos en el modelo, fueron categorizados según su tipo de convivencia escolar.

Por medio del modelo de Regresión Logística se clasificó los 6.820 RBD, obteniendo 4.813 etiquetados con “buena” convivencia (70,6%) y 2.007 etiquetados con “débil” convivencia escolar (29,4%). La distribución regional de los resultados se observa en la Tabla N°16:

Tabla N°16. Porcentaje de escuelas según tipo de convivencia y región.

Región		Buena convivencia	Débil convivencia	Total
Región de Antofagasta	Recuento	70	46	116
	% dentro de región	60,3%	39,7%	100,0%
Región de Arica y Parinacota	Recuento	54	21	75
	% dentro de región	72,0%	28,0%	100,0%
Región de Atacama	Recuento	58	34	92
	% dentro de región	63,0%	37,0%	100,0%
Región de Aysén del General	Recuento	38	19	57
	% dentro de región	66,7%	33,3%	100,0%
Región de Coquimbo	Recuento	354	89	443
	% dentro de región	79,9%	20,1%	100,0%
Región de La Araucanía	Recuento	678	143	821
	% dentro de región	82,6%	17,4%	100,0%
Región de Los Lagos	Recuento	516	129	645
	% dentro de región	80,0%	20,0%	100,0%
Región de Los Ríos	Recuento	268	44	312
	% dentro de región	85,9%	14,1%	100,0%
Región de Magallanes y de la	Recuento	33	17	50
	% dentro de región	66,0%	34,0%	100,0%
Región de Ñuble	Recuento	254	44	298
	% dentro de región	85,2%	14,8%	100,0%
Región de Tarapacá	Recuento	68	39	107
	% dentro de región	63,6%	36,4%	100,0%
Región de Valparaíso	Recuento	446	244	690
	% dentro de región	64,6%	35,4%	100,0%

		Buena convivencia	Débil convivencia	Total
Región del Biobío	Recuento	502	155	657
	% dentro de región	76,4%	23,6%	100,0%
Región del Libertador Bernardo	Recuento	319	104	423
	% dentro de región	75,4%	24,6%	100,0%
Región del Maule	Recuento	457	86	543
	% dentro de región	84,2%	15,8%	100,0%
Región Metropolitana de Sant	Recuento	698	793	1491
	% dentro de región	46,8%	53,2%	100,0%
Total	Recuento	4813	2007	6820
	% dentro de región	70,6%	29,4%	100,0%

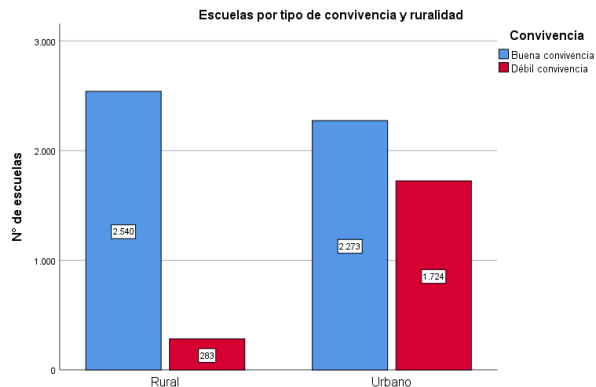
Fuente: Elaboración propia.

Como muestra la tabla, en todas las regiones se observa un mayor porcentaje de escuelas con “buena” convivencia, exceptuando a la Región Metropolitana, donde 793 escuelas (equivalentes a un 53,2% del total regional) es calificada con “débil” convivencia, mientras 698 (equivalentes a un 46,8% del total regional) se califica con “buena” convivencia. Así, del total nacional de escuelas con convivencia “débil”, alrededor del 40% se encontraría en la Región Metropolitana.

Las cinco regiones con los mayores porcentajes de escuelas con “débil” convivencia son la Región Metropolitana (53,2%), la Región de Antofagasta (39,7%), la Región de Atacama (37%), la Región de Tarapacá (36,4%) y la Región de Valparaíso (35,4%). En cambio, las cinco regiones con menores porcentajes son la Región de los Ríos (14,1%), la Región de Ñuble (14,8%), la Región del Maule (15,8%), la Región de la Araucanía (17,4%) y la Región de Los Lagos (20%).

De esta forma, se observa que las regiones del norte del país presentan porcentajes altos de convivencia “débil”, al contrario que las regiones del sur. Si agrupamos según macrozona norte, centro y sur, sin considerar a la Región Metropolitana debido a la gran cantidad de escuelas que posee, la macrozona norte muestra un 27,5% (229) de sus escuelas con “débil” convivencia escolar, la macrozona centro un 24,2% (633) y la macrozona sur un 18,7% (352).

En cuanto a las escuelas rurales, se observa que un 90% (2540) está calificada con “buena” convivencia mientras que un 10% con “débil” convivencia (283). Por su parte, el 56,9% (2273) de las escuelas urbanas fue catalogada con “buena” convivencia mientras que un 43,1% (1724) fue catalogada con convivencia “débil”. Gráficamente las frecuencias son:

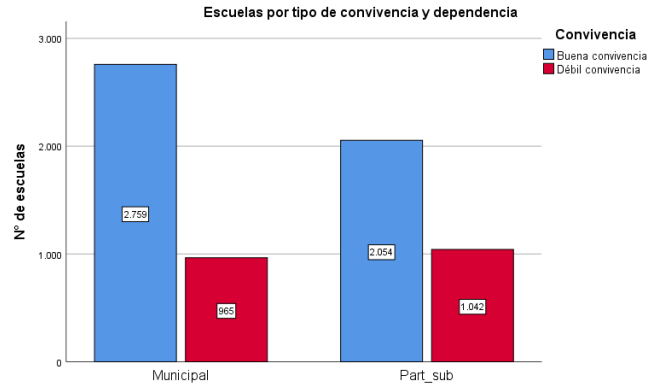


Al observar las escuelas rurales según macrozona¹⁶, se constata que en el norte un 92% (310) presenta “buena” convivencia, mientras que un 8% (27) muestra convivencia “débil”. En la macrozona centro, un 91,3% (1022) presenta “buena” convivencia versus un 8,8% (98) de escuelas con convivencia “débil”. Por último, en la macrozona sur, un 90% (1120) de las escuelas rurales muestra “buena” convivencia versus un 10% (125) convivencia “débil”. En consecuencia, y en proporción, la macrozona norte muestra los porcentajes más altos de “buena” convivencia en establecimientos educacionales rurales en comparación con las macrozonas centro y sur.

Es así que el 88,2% (202) de las escuelas con “débil” convivencia en la macrozona norte corresponde a escuelas urbanas, 84,51% (535) de las escuelas con “débil” convivencia en la macrozona centro corresponde a establecimientos urbanos y 64,4% (227) de las escuelas con “débil” convivencia, en la macrozona sur, corresponde también a zonas urbanas.

Respecto a la dependencia, un 74,1% (2759) de escuelas municipales presentan “buena” convivencia, mientras que 25,9% (965) “débil” convivencia. En el caso de las escuelas particulares subvencionadas, un 66,3% (2054) exhiben convivencia “buena” mientras que 29,4% (1042) muestran convivencia débil. Gráficamente los resultados son:

¹⁶ Sin considerar a la Región Metropolitana.



Consistentemente con la teoría y considerando la importancia de las variables para el modelo, se observa que los establecimientos con “buena” convivencia presentan menores cifras en la variable *cest_p16_02* (algún compañero te ha pegado), cifras superiores en la variable *cest_p12_03* (“mis profesores explican por qué es importante cumplir las normas”) y en la variable *cest_p05_01* (“mis profesores me dicen que soy buen estudiante”), etc.

5 Limitaciones, conclusiones y trabajo futuro

5.1 Limitaciones

López et al. (2018) han planteado la existencia de un escenario ambivalente para la implementación de políticas de convivencia escolar en el país. Como una consecuencia de aquella situación, han recomendado revisar, y posiblemente modificar, la forma en que está definido, cómo es medido y cómo es reportado el Indicador de Clima de Convivencia Escolar de la Agencia de Calidad de la Educación. Los mismos autores han llamado a revisar el sistema de rendición de cuentas asociado al ICCE (López, Ascorra, Bilbao & Carrasco, 2019).

Dicha situación ofrece una explicación posible para los ajustes alcanzados por los modelos, pues, aunque los Cuestionarios de Calidad y Contexto y la Encuesta PACES buscan aproximarse a la convivencia en los establecimientos educacionales, lo hacen desde un soporte conceptual diferente¹⁷. La mirada de Programa PACES-PUCV está construida desde un paradigma de convivencia “como preparación para la democracia” y un modelo social-ecológico, mientras que la mirada de la ACE se aproxima a una perspectiva enfocada en la seguridad.

Estas diferencias se expresan en las dimensiones medidas por cada encuesta. La mirada de la ACE enfatiza en su medición aspectos disciplinarios de la convivencia, los cuales quedan en evidencia en las dimensiones que mide: ambiente de respeto, seguro y organizado. En cambio, la propuesta del Programa PACES-PUCV aborda una serie de variables que van desde la satisfacción de los estudiantes con la escuela y el apoyo docente, a su percepción del ambiente físico de las aulas y el contexto en que se inserta el establecimiento educacional. Adicionalmente, la ACE excluye de su medición dimensiones que son esenciales para la convivencia escolar en la mirada del Programa PACES-PUCV. Es así como la Encuesta PACES incluye dentro de su medición la participación de los estudiantes, mientras que la ACE mide de forma separada a través de los IDPS el “clima de convivencia escolar” y “participación y formación ciudadana”¹⁸.

Pese a lo anterior, es importante destacar que en el concepto de “participación” utilizado por la ACE predominan claramente relaciones jerárquicas, en que el estudiante actúa dentro de las reglas generadas por los adultos y no es co-creador de éstas, a través de un proceso deliberativo

¹⁷ Es así que, el concepto de clima de convivencia escolar subyacente al indicador IDPS ICCE se orienta al orden y la disciplina dentro de las comunidades educativas, buscando limitar la violencia y generar espacios seguros como un medio la mejora en la calidad de los aprendizajes. Esto queda en evidencia al observar, por ejemplo, la relación jerárquica profesor-alumno que están a la base los Cuestionarios de Calidad y Contexto de la Educación.

¹⁸ Sin embargo, la misma separación del IDPS ICCE y IDPS de participación y formación ciudadana puede ser interpretada “[...] como un sinsentido, porque convivir implica participar, siendo la participación un señalamiento de justicia en una sociedad creada y construida por unos pocos” (Ascorra y López, 2019: 143).

de los actores en las comunidades educativas, como exigiría una convivencia de orientación democrática. Es relevante, entonces, desde que perspectiva se comprende la convivencia.

Todo lo anterior sugiere que la principal limitación de este trabajo fue la baja disponibilidad de variables o ausencia de datos administrativos suficientes para emular las dimensiones consideradas por la Encuesta PACES, lo cual se añade a los diferentes paradigmas subyacentes a los instrumentos de medición de convivencia escolar en el país¹⁹.

Además, las variables 2018 “sin equivalencia” en 2016 no fueron incluidas en la base de datos del estudio (aunque apuntaban directamente a las características de la convivencia en la escuela). El término “sin equivalencia” indica solamente que la variable no estuvo presente en los Cuestionarios de Calidad y Contexto del año 2016. Están en este caso `cest_p11_01` (“las normas de convivencia son conocidas por todos los estudiantes”), `cest_p11_02` (“las normas de la escuela son claras”), `cest_p11_03` (“los estudiantes respetamos las normas de convivencia”), `cest_p11_04` (“los profesores nos enseñan las consecuencias de no cumplir las normas”), `cest_p18_03` (“mis compañeros resuelven sus problemas conversando”) y `cest_p10_02` (“mis compañeros cuidan que la escuela esté limpia...”)

En otras palabras, los resultados no debiesen sorprender en la medida que diferentes variables relevantes no están presentes para incluirlas en el modelo. Por ejemplo, si nuestras variables incluyen solo las variables SIMCE, es esperable que los resultados obtenidos sean distintos al alcanzado en una estimación posible con todas las variables requeridas. La diferencia incluso podría ser más pronunciada a causa de las notorias diferencias conceptuales entre las variables.

Además, debido a la muestra de datos PACES disponibles, este estudio también se vio limitado al nivel de educación básica, y a establecimientos subvencionados y municipales, por lo cual representa una parte de la realidad del sistema educativo del país. Al no estar expuestos a los datos de establecimientos particulares pagados, por ejemplo, los modelos no aprendieron de ellos.

Un aspecto adicional para considerar es la elección del punto de corte para clasificar a los establecimientos educacionales con “débil” o “buena” convivencia escolar, que en este caso usó desviaciones estándar respecto a la media de las variables “clima de aula” y “clima escolar”. Este punto debiese equilibrar al menos dos aspectos: a) el criterio de los expertos en política educativa para clasificar establecimientos en una u otra categoría y b) que mientras mayores son las desviaciones estándar usadas para clasificar, más pequeño es el grupo de escuelas con convivencia “débil” lo cual produce un problema de clasificación con clases desbalanceadas²⁰.

¹⁹ Además, como muestra la Tabla N°6, variables relevantes no fueron incluidas en el modelo ya que no estuvieron disponibles para los años en análisis (2016 y 2018).

²⁰ Mientras mayor sea la diferencia entre clases más se adaptará a la clase mayoritaria lo que limita el desempeño predictivo del clasificador. Como señala Bagnato (2020) para abordar este problema usar las siguientes estrategias para abordar el problema: a) obtener más muestras; b) eliminar muestras del set mayoritaria; c) crear muestras sintéticas definiendo algún

Se estima que otra limitación de este estudio aplicado es el uso de clasificación binaria: “débil” o “buena” convivencia. Esto no se debe al uso de la clasificación binaria en sí misma, que puede ser suficiente en algunos problemas de clasificación, sino a su aplicación a este problema en particular. En ese sentido podría ser deseable una clasificación multiclase en niveles de convivencia o incluso una clasificación multiclase según las dimensiones de la Encuesta PACES.

Otra limitante de este estudio es que la etiqueta no es el atributo real de las escuelas, es una imputación que muestra el riesgo del establecimiento de presentar problemas de convivencia. Es una forma de saber dónde debería ir a mirar la convivencia escolar, pese a que no la hemos medido, partiendo de donde tengo la sospecha que puede ser peor. Ésta no reemplaza una medición efectiva de las variables de la Encuesta PACES en cada establecimiento.

Además, aunque los resultados permiten saber qué escuelas podrían ser parte de un grupo de riesgo por problemas de convivencia escolar, no permite saber cuáles son concretamente los problemas de convivencia a que se enfrentan. Es decir, la focalización de escuelas a través del modelo debería complementarse con la identificación de problemáticas y levantamiento de información en terreno.

Por último, es importante mencionar como limitante el tamaño de la muestra disponible de la Encuesta PACES para entrenar a los algoritmos ($n=1.346$), que fue el resultado de agrupar las respuestas de los estudiantes a nivel de establecimiento. Esta situación nos obligó a utilizar la validación cruzada en cada uno de los modelos a fin de evitar el sobreajuste.

En resumen, las limitaciones más importantes para esta investigación fueron:

1. Baja disponibilidad de variables o ausencia de datos administrativos suficientes para emular las dimensiones consideradas por la Encuesta PACES.
2. Que existen diferentes paradigmas subyacentes en los instrumentos de medición de convivencia escolar en el país.
3. Existencia de variables 2018 “sin equivalencia” el año 2016, que no pudieron ser incluidas dentro del estudio.
4. El tamaño de la muestra para entrenar los algoritmos (Encuesta PACES), la cual, al agruparse en establecimientos educacionales, se redujo considerablemente.

5.2 Conclusiones

El presente documento abordó la imputación de una medición de la convivencia escolar a escuelas que no contaban con ella, específicamente 6.820 establecimientos educacionales de educación básica 2018. Esta medición, tiene atributos que conviene destacar como son no ser multi informante, medir dimensiones complejas de la convivencia escolar, partir de un marco

criterio; d) usar métodos de ensamble y e) penalización de la clase mayoritaria. Todas estas estrategias tendrán sus limitaciones y ventajas, pero no reemplazan contar con los datos reales.

teórico social-ecológico, recoger evidencia de experiencias internacionales y un diagnóstico holístico de la convivencia escolar.

En el primer capítulo de este documento se revisó en la literatura especializada los distintos conceptos de convivencia escolar que están presentes (objetivo específico N°1). Entre los conceptos identificados, este documento se enfoca en la mirada de la “convivencia escolar” como “preparación para la democracia” desde un enfoque de derechos, siguiendo al Programa PACES.

Mirada la convivencia como “preparación para la democracia”, y partir de la literatura, se identifican factores que estarían involucrados en las formas de convivir en las escuelas. Considerando una mirada ecológica se incluyen: a) factores individuales; b) factores familiares; c) factores escolares; d) factores contextuales y e) factores asociados a las políticas públicas.

Luego se revisaron y compararon brevemente las medidas de convivencia escolar de PACES y de la Agencia de Calidad. Al respecto se estima que la definición oficial de la Agencia de Calidad de la Educación es un concepto más bien restringido de la convivencia.. En cambio, la mirada del Programa PACES-PUCV, muestra una perspectiva diferente tanto en su sustrato teórico, como en las finalidades declaradas por su equipo, que se define como un equipo de “investigación-acción” (Ascorra & López, 2019).

La imputación se realizó por medio de modelos de “*machine learning*” para problemas de regresión y, posteriormente, de clasificación. En el caso de los modelos de regresión el coeficiente de determinación (R^2) alcanzó valores moderados en los datos de entrenamiento y prueba.

Por su parte, en el caso de los modelos de clasificación, el mejor modelo (Regresión Logística) obtuvo un *accuracy* del 78,51%, logrando una *precisión* del 66,1% y un *recall* del 81,25% para escuelas con “débil” convivencia, mostrando así capacidad para identificar y clasificar correctamente ambas categorías de escuelas y disminuyendo a la vez los “falsos negativos”.

Es importante destacar que existen diversos métodos que buscan reducir los errores de clasificación de los modelos. Para esa tarea es primordial distinguir entre errores, porque hay errores que pueden ser permisibles, mientras que otros se deben evitar a toda costa²¹. Para clasificar los errores que comete un modelo se usa la “matriz de confusión”, una herramienta visual que nos permite entender la eficiencia de un algoritmo de clasificación.

En nuestro caso, se considera permisible el error de clasificar una escuela con “buena” convivencia escolar con la etiqueta de “débil” convivencia y se busca reducir al mínimo las escuelas con “débil” convivencia clasificados con la etiqueta “buena” convivencia. Es por eso que una métrica clave para la clase de convivencia “débil” es el *recall*, ya que nos interesa

²¹ Por ejemplo, si se requiere detectar enlaces maliciosos, un error de clasificación puede llevar a etiquetar un enlace seguro como peligroso, con lo que no se corre gran riesgo. En cambio, un fallo a la hora de detectar un enlace no seguro sí supone una amenaza tangible

disminuir al mínimo los “falsos negativos”, es decir, escuelas que teniendo “débil” convivencia, son clasificados con “buena” convivencia. El modelo obtuvo un *recall* superior al 80%.

Es así que esta investigación aplicada cumplió con el objetivo “imputar el resultado del indicador de convivencia escolar PACES a aquellos establecimientos que no hayan contestado la Encuesta de Convivencia Escolar (objetivo específico N°4)” se utilizó el modelo de regresión logística para clasificar los 6.820 RBD 2018, obteniendo 4.813 etiquetados con “buena” convivencia (70,6%) y 2.007 etiquetados con “débil” convivencia escolar (29,4%).

Las variables más importantes en la clasificación fueron *cest_p16_02* (“algún compañero te ha pegado”), *cest_p12_03* (“mis profesores explican por qué es importante cumplir las normas”), *cest_p05_01* (“mis profesores me dicen que soy buen estudiante”), *cest_p21_02* (“me siento orgulloso de mi escuela”), la categoría de desempeño “insuficiente”, *cest_p09_01* (“los estudiantes de mi curso nos tratamos con respeto”), *cest_p18_05* (“mis profesores me animan a decir mis opiniones”) y *cest_p22_02* (“el director me hace sentir que soy parte importante de mi escuela”).

De la imputación, se evidencia que las regiones del norte del país presentan porcentajes altos de convivencia “débil”, al contrario que las regiones del sur. En cuanto a las escuelas rurales, se observa que un 90% (2.540) está calificada con “buena” convivencia mientras que un 10% con “débil” convivencia (283). Por su parte, el 56,9% (2.273) de las escuelas urbanas fue catalogada con “buena” convivencia mientras que un 43,1% (1.724) fue catalogada con convivencia “débil”. Destaca la situación de la Región Metropolitana, que concentra alrededor del 40% de las escuelas con convivencia “débil”.

Además, de las brechas de convivencia escolar detectadas a nivel regional, a nivel comunal también se observan brechas significativas. Al observar los porcentajes de convivencia “débil” por comuna, destacan Peñalolén (71,4%) La Pintana (69,4%), La Florida (68%), entre otras.

Considerando el detalle de los resultados de la imputación, esta herramienta podría tener los siguientes usos:

- Ser un criterio para la focalización de planes y programas sociales del Mineduc en aquellas escuelas que más necesitan de apoyo en materia de convivencia escolar.
- Identificación temprana, por parte de los sostenedores de la escuela²², de establecimientos educacionales con “débil” convivencia, los cuales presentan mayor riesgo de problemas de convivencia escolar, a fin de actuar preventivamente y ampliar la capacidad de los servicios

²² Los cuales pueden ser Municipales, Particulares Subvencionados, de Administración Delegada, o Servicios Locales de Educación.

públicos responder a las demandas de educación de calidad. Esta sería una alarma que permita reducir las escuelas en esa situación²³.

- Identificar establecimientos educacionales que, presentando dificultades de convivencia escolar, muestren riesgo de problemas sociales emparentados con la convivencia “débil”, como son la repitencia y la deserción o exclusión educativa de los estudiantes.
- Identificar establecimientos educacionales que, presentando una buena convivencia escolar, puedan servir de modelo a otras comunidades educativas para la gestión de su propia convivencia escolar. Esto permitiría, además, pasar de la independencia a la interdependencia y el trabajo colaborativo, recogiendo las experiencias de otras escuelas, y afianzando esta práctica como forma de trabajo.

Respecto al uso del modelo, es importante definir un período de actualización del cálculo, de manera que los resultados se mantengan actualizados con los últimos datos disponibles.

Finalmente, se considera fortaleza de este trabajo la realización de un estudio basado en “aprendizaje automático” en el contexto de un programa de postgrado en políticas públicas, ya que no está entre las herramientas más usadas actualmente. El valor está en que el desafío de la convivencia escolar es enorme y amerita explorarlo con las distintas herramientas disponibles hoy. La convivencia no solo es una red de relaciones en las escuelas, sino también la oportunidad de construir una sociedad más cohesionada, pacífica y justa, que contribuya a transitar hacia niveles cada vez mayores de bienestar individual y colectivo.

El contexto educativo nacional ha cambiado y, aunque la convivencia presencial se ha interrumpido, sigue ocurriendo en entornos virtuales. En este escenario, la convivencia escolar se produce a distancia, y tal vez el aparezca con más fuerza el ciberacoso como problema, en un contexto donde las relaciones escolares se han trasladado a entornos virtuales.

Los códigos de relacionamiento de la convivencia virtual pueden ser distintos a los de la convivencia presencial, pero la reducción de la violencia escolar, el *bullying* y el *ciberbullying* siguen presentes, así como la necesidad del desarrollo de un sentido de comunidad, habilidades socioemocionales, y aumento de la participación escolar.

5.3 Trabajo futuro

Bajo la hipótesis de que era posible imputar los resultados de la Encuesta PACES, este trabajo utilizó datos administrativos como insumo de clasificación. Los resultados obtenidos muestran

²³ La intervención podría partir por evitar la ausencia de planes de gestión de la convivencia e instalar la habilidad para promover prácticas cotidianas dentro de la escuela orientadas a negociar las normas y resolver los conflictos de forma democrática. Hasta abordar desde los instrumentos de gestión de los establecimientos educacionales, como el Plan de Gestión de la Convivencia Escolar (PGCE), Reglamento Interno, Proyecto Educativo Institucional (PEI), Plan de Mejoramiento Educativo (PME), entre otros

que los datos administrativos pueden contribuir a caracterizar de mejor forma la realidad educativa de los establecimientos educacionales del país.

En ese contexto, se estima necesario ahondar en las siguientes líneas de investigación:

- Avanzar en una imputación de valores continuos usando modelos de regresión, para lo cual se requiere una mayor cantidad (y calidad) de datos administrativos.
- Realizar una clasificación multiclase, que permita comprender de forma un poco más detallada el tipo de convivencia de cada establecimiento educacional.
- Indagar sobre el vínculo de escuelas con “débil” convivencia y su relación con otros problemas sociales como la repitencia y la exclusión o abandono escolar.
- También existe la posibilidad de que construir el valor de final de la convivencia del establecimiento desde los resultados individual de estudiantes, padres y profesores pudiera tener mejores resultados.
- Por último, implementar clasificación binaria y multiclase con algoritmos de aprendizaje reforzado a fin de evaluar si las métricas de resultado mejoran.

6. Bibliografía

- Agencia de Calidad de la Educación (2019). *Informe Técnico Categorías de Desempeño*. Santiago de Chile.
- Agencia de Calidad de la Educación (s.f.). *Desarrollo Personal y Social: Otros Indicadores de la Calidad Educativa*. Disponible en: <https://bit.ly/2QY1sVf>
- Agencia de Calidad de la Educación (2017). *Construyendo juntos: claves para la convivencia escolar*. Santiago de Chile.
- Agencia de Calidad de la Educación (2016). *Metodología de Construcción de los Indicadores de Desarrollo Personal y Social*. Disponible en: <https://bit.ly/38eTPAZ>
- Amat, Rodrigo (2020). *Gradient Boosting con Python*. Disponible en: <https://bit.ly/39VqEEP>
- Apple, M. (2018). *¿Puede la educación cambiar la sociedad?* Santiago: LOM ediciones.
- Armijo, M. (2011). *Planificación estratégica e indicadores de desempeño en el sector público*. Manuales 69, Naciones Unidas Comisión Económica para América Latina y el Caribe (CEPAL).
- Ascorra, P., López, V., & Urbina, C. (2016). Participación estudiantil en escuelas chilenas con buena y mala convivencia escolar. *Revista de Psicología*, 25 (2),1-18.
- Ascorra, P., López, V. (2019). *Una década de investigación en Convivencia Escolar*. Valparaíso: Ediciones Universitarias de Valparaíso.
- Ascorra, P., Vásquez, P., Passalacqua, N., Carrasco, C., López, V., Núñez C. y Álvarez, M. (2017). Discursos en torno a la administración del conflicto y su relación con la escolar. *Cultura, Educación y Sociedad* 8 (2), 21-52. Disponible en: <https://bit.ly/3iOPMjg>
- Ascorra, P., López, V., Núñez, C., Bilbao, M., & Morales, M. (2016). *Relación entre segregación y convivencia escolar en escuelas públicas chilenas con altos puntajes PISA 2009*. Disponible en: <https://bit.ly/2ZN7f2O>
- Banz, C. (2008). *Convivencia Escolar*. Convivencia Escolar. Santiago, Chile: Valoras UC.
- Bagnato, J. (2020). *Aprende machine learning en español*. España: Leanpub.
- Ball, S. (1990). Foucault y la educación: disciplinas y saber. Madrid: Fundación PAIDEIA.
- Bazdresch P. (2009). La vida cotidiana escolar en la formación valoral: un caso. REICE. Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación, 7(2), 50-71.
- Bellei, C. (2015). *El gran experimento: mercado y privatización de la educación chilena*. Santiago, LOM ediciones.
- Betancourt, D. (30 de junio de 2017). *Indicadores de gestión: Definición, elaboración e interpretación con ejemplo práctico*. Recuperado el 12 de abril de 2021, de Ingenio Empresa: www.ingenioempresa.com/indicadores-una-guia-incompleta.
- Biblioteca del Congreso Nacional (2021). Historia de la Ley N° 20.529. Disponible en: <https://bit.ly/3y0XcXu> (julio, 2021).
- BID (2019). Caja de herramientas para la visualización de datos de las oficinas nacionales de estadísticas. Disponible en: <https://bit.ly/2uTtPeJ>

- Bonnefoy, J. & Armijo, M. (2005). *Indicadores de Desempeño en el sector público*. Manuales 45, Naciones Unidas Comisión Económica para América Latina y el Caribe (CEPAL).
- Bourdieu, P. (2008). *Los herederos, los estudiantes y la cultura*. México, Siglo XXI editores.
- CEPAL (2018). *La ineficiencia de la desigualdad*. Documento disponible en: <https://bit.ly/2tadzFC>
- Carbajal, P. (2016). Convivencia democrática en las escuelas. Apuntes para una reconceptualización. *Revista iberoamericana de evaluación educativa*, 6(2). Disponible en: <https://bit.ly/3j7h58w>
- Carrasco, C., López, V., & Estay, C. (2012). Análisis Crítico de la Ley de Violencia Escolar de Chile. *Psicoperspectivas*, 11(2), 31-55. <https://dx.doi.org/10.5027/psicoperspectivas-Vol11-Issue2-fulltext-228>
- Cid, P., Díaz, A., Pérez, M., Torruella, M., y Valderrama, M. (2008). *Agresión y violencia en la escuela como factor de aprendizaje escolar*. Disponible en: <https://bit.ly/2FYsSct>
- Correa, S. Ruiz-Tagle, P. (2010). Ciudadanos en democracia: fundamentos del sistema político chileno. Editorial Debate.
- Chaux, E. (2012). *Educación, Convivencia y Agresión Escolar*. Bogotá, Universidad de los Andes.
- Cheng, Y. (1994). Classroom environment and student affective performance: an effective profile. *Journal of Experimental Education*, Spring, Vol.62 (3) 221-239.
- Davenport, T. (2016). "7 ways to introduce AI into your organization". *Harvard Business Review*. Disponible en: <https://hbr.org/2016/10/7-ways-to-introduce-ai-int...> [Consulta: 16 enero de 2020].
- Drucker, P. (1954), *The Practice of Management*, Harper & Row, Nueva York.
- De Keijzer, B. (1997). *El varón como factor de riesgo: Masculinidad, salud mental y salud reproductiva*. Disponible en: <https://bit.ly/3mJdyM>
- Cea, M. (2001). *Metodología Cuantitativa: estrategias y técnicas de investigación social*. Editorial Síntesis Sociología.
- Delors, J. (1996): "Los cuatro pilares de la educación" en *La educación encierra un tesoro*. Informe a la UNESCO de la Comisión internacional sobre la educación para el siglo XXI, Madrid, España: Santillana/UNESCO. pp. 91-103.
- Fierro, C., Carbajal, P. (2019). *Convivencia escolar: una revisión del concepto*. Disponible en: <https://bit.ly/2Fvt6Ct>
- Fernández, I. (1999). *Prevención de la violencia y resolución de conflictos*. Madrid: Narcea.
- Foucault, M. (1970). *El orden del discurso*. Buenos Aires, Siglo XXI editores.
- Focus (2018). *Estudio sobre la gestión de la convivencia escolar entre los sostenedores de los establecimientos municipales: informe final*. Disponible en: <https://bit.ly/39KIxE0>
- Freire, P. (2016). *La pedagogía del oprimido*. Santiago: editorial Indómita.
- Fuentes, L., Pérez, L. (2019). *Convivencia escolar: una mirada desde las familias*. Disponible en: <https://bit.ly/35Gkung>

Gardner, Howard (1994). Estructura de la mente: la teoría de las inteligencias múltiples. México, Fondo de la Cultura Económica.

Goldsmith, S., Crawford, S., & Weinryb Grohsgal, B. (2016). Innovaciones en la prestación de servicios públicos: Número 4: El análisis predictivo: Impulsando la mejora a partir de los datos.

Gutiérrez, C. (2012). Chile: ¿eliminar o reproducir la segregación social?: dos siglos de disputas sobre modelos educacionales. Santiago, Ceibo ediciones.

Ianni, N. (2003). *La convivencia escolar: una tarea necesaria, posible y compleja*. Disponible en: <https://bit.ly/2Evr2Kq>

López, V., Ascorra, P., & Bilbao, M. (2012). El efecto del Ambiente Escolar sobre los Resultados PISA 2009. Informe Final presentado al Fondo de Investigación y Desarrollo en Educación (FONIDE). Santiago: Ministerio de Educación.

López, V., Ascorra, P., Bilbao, M., Carrasco, C., Morales, M., Villalobos, B., & Ayala, Á. (2013). Monitorear la convivencia escolar para fortalecer (no disminuir) las capacidades de las escuelas. *Revista Iberoamericana de Evaluación Educativa*, 6(2), 201-219.

López, V., Carrasco, C., Morales, M., Ayala, A., López, J. & Karmy, M. (2011). Individualizando la violencia escolar: Análisis de prácticas discursivas en una escuela municipal de la Región de Valparaíso. *Psyche*, 20(2), 75-91.

López, V., Ascorra, P., Bilbao, M. & Carrasco, C., (2019). De la violencia a la convivencia escolar: una década de investigación. En Ascorra, P. & López, V. (Eds.). *Una década de investigación en convivencia escolar* (pp. 15-50). Valparaíso: Ediciones Universitarias de Valparaíso.

López, V., Bilbao, M., Ascorra, P., Carrasco, C., Benbenishty, R., Astor, R. & Álvarez, J., (2019). El sistema de monitoreo de la convivencia escolar PACES-PUCV: una experiencia de transferencia a la política pública. En Ascorra, P. & López, V. (Eds.). *Una década de investigación en convivencia escolar* (pp. 161-185). Valparaíso: Ediciones Universitarias de Valparaíso.

López, V. (2020). School violence and convivencia escolar in Chile. En Cohen & Espelage (Eds.), *Feeling Safe in School: Bullying and Violence Prevention Around the World* (pp. 57-74). Cambridge, Harvard Education Press.

López, V., Ramírez, L., Valdés, R., Ascorra, P., & Carrasco-Aguilar, C. (2018). Tensiones y nudos críticos en la implementación de la(s) política(s) de convivencia escolar en Chile. *Calidad en la Educación*, (48), 96-129.

Magendzo A., Toledo M., & Gutiérrez V. (2013). Descripción y análisis de la Ley sobre Violencia Escolar (N°20.536): dos paradigmas antagónicos. *Estudios pedagógicos (Valdivia)*, 39(1), 377-391. Disponible en: <https://bit.ly/3i4e08f>

Martínez-Otero, V.(2001). *Convivencia escolar: problemas y soluciones*. Disponible en: <https://bit.ly/2sBhNGj>

Medina, F., Galván, M. (2007). *Imputación de datos: teoría y práctica*. Disponible en: <https://bit.ly/2RmqwWe>

Melero, D. (2020). *Política Nacional de convivencia escolar en tiempo de pandemia*. Disponible en: <https://bit.ly/3z8vfOU>

- Moos, R., Moos, B.S. y Tricket, E. (1989). Escalas de clima social: Familia, Trabajo, Instituciones Penitenciarias, Centro Escolar. Madrid: Consulting Psychologists Press.
- Morales, M., López, V. (2019). *Políticas de Convivencia Escolar en América Latina: Cuatro Perspectivas de Comprensión y Acción*. Disponible en: <https://bit.ly/2QM2CSE>
- Milicic, N., Aron, A. (2017). *Clima Social Escolar y Desarrollo Personal*. Santiago de Chile, Ediciones UC.
- Mineduc (2019a). *Clima de Convivencia Escolar*. Disponible en: <https://bit.ly/38dVBBg>
- Mineduc (2019b). Política Nacional de Convivencia Escolar. Santiago de Chile.
- Morales, M., & López, V. (2019). Políticas de convivencia escolar en América Latina: Cuatro perspectivas de comprensión y acción. *Archivos Analíticos de Políticas Educativas*, 27 (5).
- Murphy, K. (2012). *Machine Learning: A Probabilistic Perspective*. London: The MIT Press.
- Nail, O. (2013). Análisis de incidentes críticos de aula: una herramienta para el mejoramiento de la convivencia. Ril editores.
- Palma, A. (2009). Las políticas públicas que no contribuyen a la cohesión social. Editorial Catalonia.
- Konold, T., Cornell, D., Jia, Y & Malone, M (2018). School Climate, Student Engagement, and Academic Achievement: A latent Variable, Multilevel Multi-Informant Examination.
- Paredes, V., Frazer, L. (1992). *School Climate in AISD* (Austin Independent School District), Office of Research and Evaluation, AISD, Austin, Texas.
- Perales, C., Arias, E., y Bazdresch, M. (2014). *Desarrollo socioafectivo y convivencia escolar*. México: ITESO.
- Raschka, S., Mirjalili, V. (2019). *Python Machine Learning: Aprendizaje automático y aprendizaje profundo con Python, scikit-learn y TensorFlow*. España: Editorial Marcombo.
- Ramírez, M. (2016). *Convivencia escolar y ambientes de aprendizaje en una escuela primaria de San Luis potosí*. Disponible en: <https://bit.ly/2uzAk6p>
- Ramírez, M., Ascorra, P., Carrasco, C & López, V (2019). *¿Cómo entendemos en Chile una buena convivencia?* Disponible en: <https://bit.ly/2ZekpbZ>
- Redón, S. (2011). Escuela e identidad: Un desafío docente para la cohesión social. *Polis*, 10(30), 447-476. Disponible en: <https://bit.ly/360cLmq>
- Robinson, K. (2018). *Tú, tu hijo y la escuela: el camino para darle la mejor educación*. Argentina, Editorial Grijalbo.
- Rodríguez, P., Suchan, K., Truffello, R & Matas, M. (2016). Apoyando la formulación de políticas públicas y toma de decisiones en educación utilizando técnicas de análisis de datos masivos: el caso de Chile.
- Rodríguez, P., Palomino, N., & Mondaca, J. (2017). El uso de datos masivos y sus técnicas analíticas para el diseño e implementación de políticas públicas en Latinoamérica y el Caribe (Discussion Papers & Presentations). Inter-American Development Bank.
- Rosati, G. (2019). Imputación de datos perdidos mediante técnicas de Machine Learning: un experimento usando la Encuesta Permanente de Hogares: <https://bit.ly/3OURN5b>

Hernández Sampieri, R., Fernández, C. y Baptista, P. (2010). *Metodología de la investigación* (5° Ed.). México, D.F., México: McGraw Hill Interamericana.

Schiavo-Campo, S. & Tommasi, D. (1999). *Managing Government Expenditure*. Asian Development Bank, Manila.

Schmitz, H. Sikkink, K. "International Human Rights". En *Handbook of International Relations*". Edited by W. Carlsnaes; T. Risse y B. Simmons. London: Sage.

UNICEF (1989). *Convención sobre los Derechos del Niño*. Disponible en: <https://bit.ly/2ELNYoQ>

Schmarzo, B. (2013). *Big Data: Understanding How Data Powers Big Business*. Wiley.

Serrano, J., Pons, R. (2011). *El constructivismo hoy: enfoques constructivistas en educación*. Disponible en: <https://bit.ly/2ThF9LT>

Siegel, E. (2013). *Analítica predictiva: predecir el futuro usando Big Data*. Madrid, Anaya Multimedia.

Subirats, J. & Alegre, M. (2006). Convivencia social y convivencia escolar. Cuadernos de Pedagogía, 359, 12-18.

Supereduc (2020). *Informe de Resultados: Encuesta de Participación Ciudadana*. Disponible en: <https://bit.ly/2PZXnCc>

Tomar, L., Guicheney, W., Kyarisiima, H., & Zimani, T. (2016). *Big Data in the public sector: Selected applications and lessons learned*

Tuvilla, J. (2004). *Convivencia escolar y resolución pacífica de conflictos*. Sevilla. Consejería de Educación y Ciencia.

Umaña, M (2020). Clima de convivencia escolar en el marco de la complejidad. *Contextos: Estudios De Humanidades Y Ciencias Sociales*, (47). Recuperado a partir de <http://revistas.umce.cl/index.php/contextos/article/view/1566>.

UNESCO (2008). Segundo Estudio Regional Comparativo y Explicativo (SERCE). Informe Final.

Waissbluth, M. (2018). *Educación para el siglo XXI: el desafío latinoamericano*. Santiago, Editorial Fondo de Cultura Económica.

Waissbluth, M. (2010). *Se acabó el recreo: la desigualdad en la educación*. Santiago, Editorial Debate.

López, V., Díaz, H., & Carrasco, C. (2015). *Nosotros sí podemos: Aprendiendo a mejorar la convivencia escolar*. Valparaíso, Chile: Pontificia Universidad Católica de Valparaíso y Centro de Investigación Avanzada en Educación (CIAE).

Zaitegui, N. (2020). *La Educación en y para la Convivencia Positiva en España*. Disponible en: <https://bit.ly/319w6rM>

Videografía

Educación para la convivencia y resolución de conflictos. Nélica Zaitegi, pedagoga: <https://www.youtube.com/watch?v=nNODDtwY1uM>

VIDEO FORMAS DE IMPUTACIÓN DE DATOS PÉRDIDOS QUE ENTREGA SPSS:
<https://www.youtube.com/watch?v=nwCIBKS1uIA>

PACES (2018). *Seminario PACES-PUCV 10 años*. Disponible en:
<https://www.youtube.com/watch?v=QSGqXM1evM4>

Supereduc (2020). *Convivencia escolar y habilidades socioemocionales en pandemia*.
Disponible en: <https://www.youtube.com/watch?v=6pny09NCa64>