

TÉCNICAS DE  
*MACHINE-LEARNING*  
UTILIZADAS EN  
ESTUDIOS  
PSICOLÓGICOS CON  
ADOLESCENTES:  
UNA REVISIÓN  
SISTEMÁTICA

MACHINE-LEARNING  
TECHNIQUES USED IN  
PSYCHOLOGICAL STUDIES  
WITH ADOLESCENTS: A  
SYSTEMATIC REVIEW



Alejandro J. Mena

ORCID: 0000-0001-7502-7839

Nicolás de Oliveira Cardoso

ORCID: 0000-0002-1555-1409

Carlos Eduardo Xavier

ORCID: 0000-0001-8556-6756

Irani de Lima Argimon

ORCID: 0000-0003-4984-0345

*Pontificia Universidade Católica do Rio Grande do  
Sul, PUCRS*

email: [jandromb@hotmail.com](mailto:jandromb@hotmail.com)

**RESUMEN**

*Durante la última década ha habido un crecimiento exponencial de modelos de inteligencia artificial aplicados a los distintos campos del conocimiento incluida la psicología. Diversos estudios han utilizado dichos modelos con el objetivo de identificar precozmente potenciales riesgos. Sin embargo, pocos estudios enfocados en adolescentes han utilizado dichas técnicas. La presente revisión sistemática utilizó los*

**ABSTRACT**

During the last decade there has been an exponential growth of artificial intelligence models applied to the different fields of knowledge including psychology. Various studies have used these models in order to identify potential risks early. However, few studies focused on adolescents have used these techniques. This systematic review used

*pasos sugeridos por el modelo Prisma para identificar estudios que aplicaron técnicas de Aprendizaje automático para identificar rasgos comportamentales en adolescentes. Al aplicar los criterios de inclusión y exclusión fueron identificados 5 estudios en las bases de datos PsycNET, PubMed, Scopus, Scielo, Web of Science y Science Direct. Los principales resultados muestran que los algoritmos de Machine Learning (ML) o Aprendizaje automático principalmente utilizados de forma individual o combinada fueron regresión logística (n=4) y Support Vector Machine (SVM) (n=3) además de otros como Adaboost (n=1), Random Forest (n=1), Artificial Neural Network (ANN) (n=1) y Extreme gradient boosting (XGB) (n=1). Esta revisión destaca que la utilización de métodos de aprendizaje automático provee de herramientas predictivas confiables tanto o más que los métodos estadísticos tradicionales. Por último, la presente revisión destaca la falta de estudios que utilicen dichas herramientas en el campo de la psicología principalmente en adolescentes.*

#### **PALABRAS CLAVE**

*Machine Learning; Aprendizaje Automático; Inteligencia Artificial; Adolescentes; Revisión Sistemática.*

the steps suggested by the Prisma model to identify studies that applied Machine Learning (ML) techniques to identify behavioral traits in adolescents. When applying the inclusion and exclusion criteria, 5 studies were identified in the *PsycNET, PubMed, Scopus, Scielo, Web of Science and Science Direct databases*. The main results show that the machine learning algorithms mainly used individually or in combination, were *logistic regression (n = 4) and Support Vector Machine (SVM) (n = 3)* in addition to others such as *Adaboost (n = 1), Random Forest (n = 1), Artificial Neural Network (ANN) (n = 1), and Extreme gradient boosting (XGB) (n = 1)*. This review highlights that the use of machine learning methods provides reliable predictive tools as much or even more than the traditional statistical methods. Finally, the present review highlights the lack of studies using these tools in the field of psychology, mainly in adolescents.

#### **KEYWORDS**

Machine Learning; Artificial Intelligence; Adolescents; Systematic Review.

## **INTRODUCCIÓN**

La adolescencia es una de las etapas del ciclo vital que ha captado mayor interés de los investigadores en épocas recientes. Igualmente, la adolescencia es una importante ventana para la adaptación individual, el desarrollo y la búsqueda de

oportunidades; sin embargo, es también durante esta etapa que suelen manifestarse diversos trastornos mentales. De esa forma, la adolescencia es considerada una de las fases de mayor complejidad en el proceso evolutivo del ser humano. Además, dicha etapa se caracteriza por estar acompañada de diversas modificaciones tanto biológicas, como conductuales y sociales. Muchas de las problemáticas asociadas con dicha etapa del desarrollo, están asociadas con el manejo de emociones y regulación emocional, las cuales son comúnmente con asociadas con la asunción de riesgos, violencia y conductas antisociales (Gálvez-Nieto, Vera-Bachmann, Trizano-Hermosilla, Polanco & Salvo, 2018; Paus, Keshavan & Giedd, 2008).

Se sabe que aproximadamente el 20% de los adolescentes en todo el mundo experimentan problemas relacionados con la salud mental que incluyen depresión, trastornos de ansiedad, consumo de sustancias, violencia y suicidio (McLoughlin, Gould & Malone, 2015; Patel, Flisher, Hetrick, & McGorry, 2007). Además, la literatura señala que si estos trastornos no se identifican de manera prematura y se abordan de manera efectiva, pueden potencialmente volverse crónicos y sus consecuencias en la edad adulta son considerables y en algunos casos, difícilmente tratables (Ahulu, Gyasi-Gyamerah & Anum, 2020; Cummings, Caporino & Kendall, 2014; McLeod, Horwood & Fergusson, 2016; Pine, Cohen, Gurley, Brook & Ma, 1998; Essau, Lewinsohn, Lim, Moon-ho & Rohde, 2018; Volkaert, Wante, Van Beveren, *et al.* 2020).

Así, los primeros años de vida son críticos en el desarrollo y consolidación de enfermedades y trastornos psicológicos, por lo que identificar de manera precoz potenciales dificultades es beneficioso para el desarrollo pleno de las personas (McCory *et al.*, 2015). En el contexto de la investigación en psicología, diversos abordajes estadísticos han buscado ofrecer algún tipo de respuesta estas necesidades. Paralelamente, con el avance de la ciencia y específicamente de las ciencias computacionales, el Aprendizaje Automático ha experimentado un crecimiento acelerado, considerando que su aplicación abarca diversas áreas del conocimiento humano y permite el análisis de grandes volúmenes de datos de forma rápida. Sin embargo, en el campo de la investigación en psicología su uso es aún exiguo, a pesar de mostrar diversas ventajas y aplicaciones en diversos contextos de la investigación en psicología, como por ejemplo detectar condiciones de riesgo precozmente, prestar diagnósticos precisos y predecir tratamientos certeros (Correll, *et al.*, 2018; Dwyer & Koutsouleris, 2022; Kim, Sharma, & Ryan, 2015; Marin, 2016; Morrow, Campos Vega, Zhao, *et al.*, 2020; Sansone, 2019). Lo anterior se torna aún más evidente al observar que dicha carencia es aún más evidente en los estudios realizados con niños y jóvenes (Dwyer & Koutsouleris, 2022).

Aprendizaje automático o *Machine Learning* (ML) se define como un conjunto de métodos propios de la inteligencia artificial, los cuales permiten a las máquinas (algoritmos) aprender con base en la experiencia. De esa manera los algoritmos aprenden sin haber sido previamente programados. En ese sentido, los métodos de Aprendizaje automático permiten identificar predictores utilizando un número limitado de variables y que además suelen ser sutiles y poco identificables para otros métodos estadísticos (Luca, Kleinberg, & Mullainathan, 2016; Morrow, Campos Vega, Zhao, *et al.*, 2020; Sansone, 2019). Cabe resaltar que algunos autores resaltan la existencia de dos grandes tradiciones en el campo de la estadística, por un lado la estotástica o explicativa y por otro la algorítmica o predictiva, de manera que el aprendizaje automático se enmarca dentro de la cultura algorítmica o predictiva (Breiman, 2001; Boulesteix & Schmid, 2014). Cabe resaltar el elemento clave en la diferenciación de ambos modelos, de esa manera en el enfoque algorítmico un modelo aprende de los ejemplos en lugar de ser previamente programados por reglas, como sucede en los modelos estadísticos explicativos (Rajula, *et al.*, 2020).

En el tiempo actual, la gran cantidad de datos disponibles (accesible gracias a un conjunto de plataformas) puede ser un valioso recurso para los investigadores en psicología, tomando en cuenta que eso permite la expansión de los estudios de laboratorio y métodos tradicionales de investigación a otros contextos de la interacción humana, sin perder poder estadístico e incluso mejorando su validez externa (Sagi, 2019).

De este modo el presente estudio busca en primer lugar, resaltar la importancia y el evidente crecimiento de modelos predictivos aplicados en el campo de la psicología, particularmente con jóvenes y niños. Además, tomando en cuenta que según el conocimiento de los autores hasta la presente fecha no existe otro tipo de estudios de revisión sistemática que hayan analizado dichos abordajes de aprendizaje automático aplicados en el campo de la psicología y concretamente con adolescentes.

## MÉTODO

Para la realización del presente estudio, se optó por la elaboración de una revisión sistemática de literatura siguiendo el modelo PRISMA (Preferred Reporting Items for Systematic Review and Meta-Analyses; Moher, Liberati, Tetzlaff, & Altman, 2009).

### *Recolección de datos*

Inicialmente se hizo uso de los diccionarios de términos "Mesh" (*PubMed*) y "Thesaurus" (*PsycNet*) para dar mayor sensibilidad a la búsqueda. La Recolección de datos de datos se realizó por medio de una búsqueda online para la selección de las principales producciones científicas nacionales e internacionales utilizando los siguientes descriptores: "Machine Learning" OR "deep learning" AND "Adolescent behavior" OR "adolescent behaviors" OR "adolescent development". Las bases de datos consultadas en la presente revisión fueron: *PsycNET*, *US National Library of Medicine and National Institutes of Health (PubMed)*, *Scopus*, *Scientific Electronic Library On-line (Scielo)*, *Web of Science* y *Science Direct*. La búsqueda de los artículos se realizó en noviembre de 2019 y fueron analizados por tres jueces independientes.

### ***Crterios para la selección de los artículos***

Se tomó en cuenta como criterios de elegibilidad que fuesen estudios que analizaron rasgos de conducta o emocionales en adolescentes y que al mismo tiempo dichos rasgos fueran analizados utilizando métodos predictivos de Aprendizaje automático o inteligencia artificial. Con este fin, existe evidencia que, a pesar del avance en la identificación de condiciones psicológicas y emocionales, existe aún una evidente carencia en identificar rasgos predictivos para dichas conductas (Franklin et al. 2017).

En la presente revisión se tomó como criterio de exclusión documentos como libros, tesis, artículos de opinión, reseñas, revisiones sistemáticas, revisiones de literatura, comentarios, anales e informes. Además, se excluyó artículos que a pesar de haber evaluado adolescentes aplicando modelos de Aprendizaje automático lo hicieron con condiciones físicas o psiquiátricas como autismo, daño cerebral y enfermedades mentales. A lo largo del proceso se excluyó inicialmente artículos que no mostrasen en el título alguna de las palabras clave previamente establecidas para la búsqueda. Posteriormente los tres jueces evaluadores realizaron una lectura pormenorizada de los resúmenes, buscando identificar estudios que a pesar de incluir las palabras clave en el título no coincidían con los objetivos propuestos en la presente revisión o en algunos casos no se enmarcaban en la metodología que la presente investigación buscó identificar.

## **RESULTADOS**

La búsqueda preliminar identificó 231 estudios potencialmente útiles. Fueron encontrados 2 en la base *PsycNET*, 129 en la base *Pubmed*, 9 en *Scopus*, 88 en *Web of Science*, 2 en *Scielo* y 1 en *Science Direct*. Luego del proceso de análisis y lectura

detallada fueron incluidos 5 estudios. La figura 1 muestra un flujograma que detalla el proceso de identificación y selección de los artículos analizados.

La tabla 1 (Anexo 1) muestra una síntesis de las principales características de los estudios seleccionados y analizados. Para tener una mejor referencia y ubicarlos con mayor facilidad a lo largo del manuscrito, los mismos fueron clasificados con números ordinales (1,2,3, etc.). En tal sentido, después de una síntesis detallada de los 176 artículos encontrados, solamente 5 estudios respondieron a los criterios de inclusión planteados en la presente revisión.

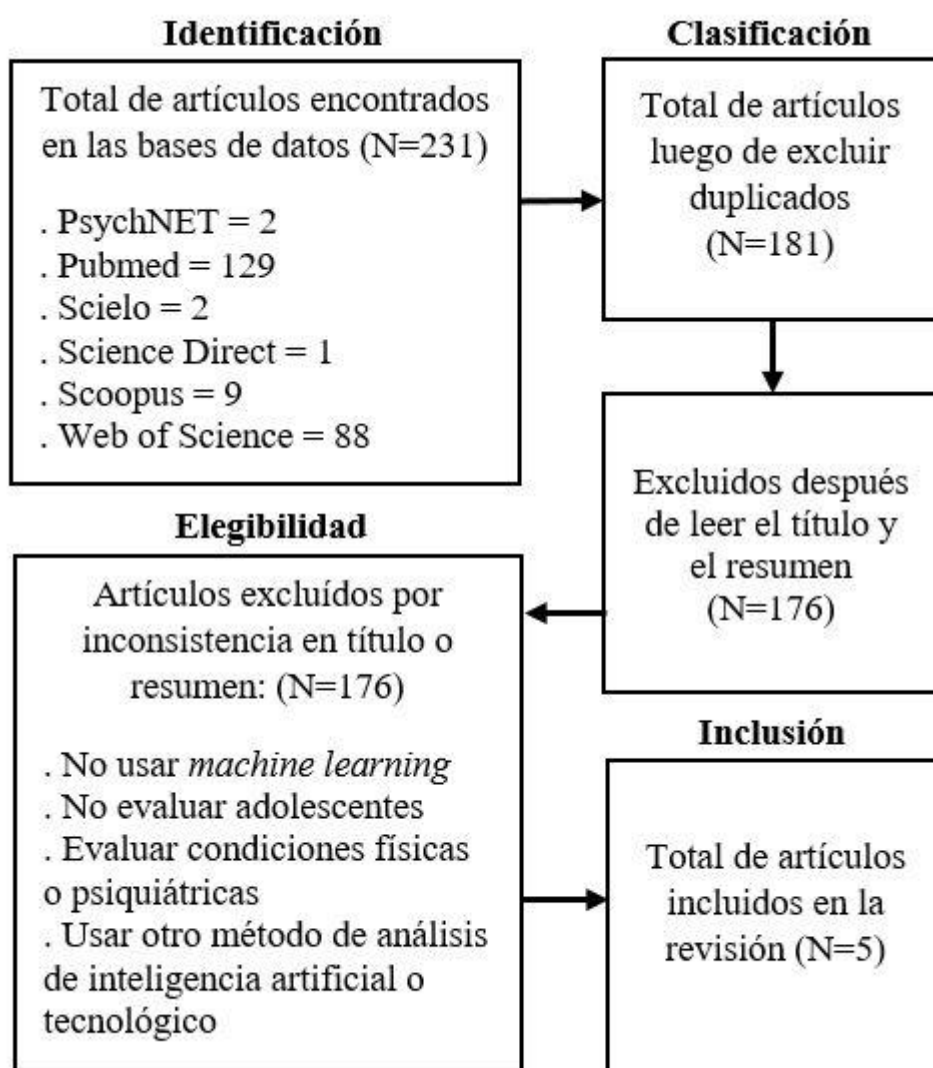


Figura 1. Flujograma de la sistematización de la revisión

Es importante destacar que de los 5 estudios analizados 3 estudios (1,3 y 5) realizaron resaltaron abiertamente una comparación entre métodos estadísticos tradicionales o análisis clínicos con abordajes de Aprendizaje automático, en todos los casos los autores resaltan el alto nivel de confiabilidad que las técnicas de aprendizaje automático mostraron por encima de los métodos estadísticos o clínicos habituales.

Por otro lado, aunque la literatura muestra que existen diversos algoritmos comúnmente utilizados en estudios de investigación en psicología, pudo notarse que el conjunto de algoritmos de aprendizaje supervisado denominado *regresión logística* (n=4), que, aunque es conocida como una técnica estadística tradicional, es al mismo tiempo uno de los algoritmos de aprendizaje automático de naturaleza binaria más utilizados para la clasificación de datos. Además, es uno de los algoritmos de aprendizaje automático más simples, el cual por medio de un enfoque matemático realiza análisis predictivo (Maulud & Abdulazeez, 2020). Por su parte el segundo conjunto de algoritmos más utilizados fue el *Support Vector Machine SVM* (n=3), el cual probablemente fue utilizado en diversos estudios por su capacidad para resolver tareas de clasificación a partir de una serie de datos o muestras que posteriormente generará nuevos resultados o muestras.

Otros algoritmos utilizados en los estudios de la presente revisión fueron *Adaboost* (n=1), Bosques aleatorios o *Random Forest* (n=1), *Artificial Neural Network ANN* (n=1) y *Extreme gradient boosting XGB* (n=1). Dichos algoritmos fueron utilizados como complemento de los algoritmos antes mencionados y los cuales además tienen la característica de ser modelos predictivos de clasificación estadística escalonada. Aunque dichos algoritmos comparten características comunes, la selección de cada uno de ellos depende de los criterios de los autores, los cuales generalmente están condicionados por el tipo de datos recolectados. Por otro lado, cabe resaltar que todos los estudios incluidos en la presente revisión incluyeron métodos de validación (Validación cruzada 3-k, 5-k y 10-k y métodos de retención).

Siguiendo con los aspectos metodológicos, un elemento que llama la atención es el hecho que, de todos los estudios analizados, únicamente uno (estudio 2) hizo uso de grupo experimental y grupo control, mientras que los restantes cuatro estudios no lo hicieron. El único estudio que utilizó grupo experimental y control tuvo como objetivo identificar posible riesgo suicida a través de la identificación de patrones verbales y no verbales tanto en jóvenes con tendencia suicida (dentro de los que se seleccionó a jóvenes con un solo intento y jóvenes con más de un intento) como en jóvenes sin dicha tendencia (Venek, Scherer, Morency & Pestian, 2017).

Respecto del contexto donde se realizó los estudios pudo identificarse que todos fueron realizados en países del denominado primero mundo, siendo los Estados Unidos de América el país con mayor número de publicaciones (n=3) país que cuenta además con una coparticipación en el estudio realizado a nivel nacional en Irlanda (n=1), finalmente el tercer país y quinto estudio analizado fue realizado por investigadores de Corea del Sur (n=1).

Retomando elementos metodológicos observados en los estudios analizados, pudo observarse que sólo el estudio de Venek, Scherer, Morency & Pestian (2017) utilizó una muestra menor de cien participantes, específicamente dicho estudio analizó a 30 estudiantes en el grupo experimental y 30 en el grupo control. Puede observarse además que la mayoría de los estudios analizó muestras extensas de participantes, considerando que dos de los cinco estudios investigó muestras nacionales de adolescentes, por ejemplo, el estudio 4 de Fitzgerald, Mac Giollabhui, Dolphin, Whelan & Dooley (2018) en Irlanda analizó 6,062 adolescentes que conformaron una muestra de las 732 escuelas de secundaria del territorio Irlandés. Por su parte el estudio 5 realizado en Korea del Sur por Jung, Park, Kim, Na, Kim, & Kim (2019) fue el estudio con la mayor muestra de participantes analizados con 59,984 adolescentes que participaron respondiendo a la encuesta nacional de conductas de riesgo para jóvenes.

Otra característica que vale la pena destacar es que todos los estudios analizaron muestras mixtas respecto del sexo de los participantes. Sin embargo aunque de manera marginal la mayoría de estudios analizó muestras predominantemente de mujeres (Barzman, Ni, Griffey, Bachtel, Lin, Jackson, ... & DelBello, 2018; Bi, Sun, Wu, Tennen & Areli, 2013; Jung, Park, Kim, Na, Kim & Kim, 2019), un estudio analizó una muestra predominantemente de hombres (Fitzgerald, Mac Giollabhui, Dolphin, Whelan & Dooley, 2018), mientras que un estudio analizó el mismo número de participantes hombres y mujeres (Venek, Scherer, Morency & Pestian, 2017). Es oportuno destacar que se incluyó el estudio de Bi, Sun, Wu, Tennen & Areli (2013) a pesar de haber analizado una muestra de estudiantes de los años iniciales de universidad, considerando que según la Organización Mundial de la Salud (2014) la adolescencia comprende el período de desarrollo humano que marca la transición entre la niñez y la vida adulta y que suele observarse entre los 10 y 19 años de edad. Finalmente, a pesar del avance de la inteligencia artificial y Aprendizaje automático, se hace evidente la falta de estudios que utilicen dichas herramientas en el campo de la psicología, principalmente en etapas iniciales del ciclo vital como es la adolescencia.

Por otro lado, más allá de los elementos metodológicos identificados en la presente revisión, también fue posible identificar una serie de patrones temáticos en



dichos estudios, lo anterior a pesar de la evidente limitación de contar con únicamente cinco estudios.

## DISCUSIÓN

El objetivo central del presente estudio fue identificar estudios que utilizaran técnicas o métodos de aprendizaje automático para la identificación precoz de diversas condiciones psicológicas relacionadas con adolescentes. En tal sentido, fue posible identificar elementos comunes dentro de los estudios analizados en la presente revisión. Así, la presente discusión será presentada de acuerdo con las temáticas identificadas para una mejor organización de la información obtenida.

### *Artículos que analizaron consumo de sustancias de manera prospectiva*

De acuerdo con los datos presentados en la tabla 1 dos de los estudios analizaron características comportamentales y hábitos de consumo de alcohol, cannabis y otras drogas en adolescentes, siendo estos el estudio 1 y el 4 (Bi, Sun, Wu, Tennen & Areli, 2013; Fitzgerald, Mac Giollabhui, Dolphin, Whelan & Dooley, 2018). Desde ese punto de vista, el estudio 1 mostró como resultados que dentro de los factores que predicen consumo excesivo en adolescentes hombres se encuentra el deseo de sentirse más competentes, así como el hecho de competir por beber más. Para los adolescentes hombres que beben moderadamente se observó que uno de los principales motivadores para beber es disminuir la ansiedad social. Por su parte para las adolescentes mujeres que reportaron altos niveles de consumo se observó que la ansiedad, la ansiedad social y la búsqueda de nuevas sensaciones fueron los principales incentivos para consumir altas cantidades de alcohol. Por su parte el estudio de Fitzgerald, Mac Giollabhui, Dolphin, Whelan & Dooley (2018) resalta que para el consumo de alcohol intervienen múltiples variables predictoras como la red de amigos, familia disintegrada, vivir en área rural, experimentar rabia y la ruptura de relacionamientos amorosos. Mientras tanto, el consumo de tabaco fue explicado por el pobre desempeño académico como principal variable predictora. Finalmente, para el consumo de cannabis las variables sexo masculino, vivir en área urbana y pobre manejo de la ira se perfilaron como las principales variables predictoras. De esa manera estos estudios corroboran algunos elementos mostrados en previas investigaciones las cuales demostraron que el consumo temprano de alcohol y drogas suele ser un importante predictor del consumo abusivo de sustancias durante la vida adulta. Igualmente, coincide y robustece lo dicho por otros autores quienes resaltan la importancia que algunos rasgos comportamentales tienen en el inicio del consumo de sustancias como el pobre manejo de emociones, la presión social y dificultades de

tipo social o familiar (Montes, Witkiewitz, Pearson y Leventhal, 2019; Riley, Rukavina & Smith, 2016).

### ***Artículos que analizaron potenciales conductas suicidas de manera prospectiva***

Otra categoría observada dentro de los estudios detallados está relacionada con otra conducta de riesgo, en este caso ideación o tentativas suicidas en adolescentes. Así, las investigaciones dos y cinco analizaron dicho fenómeno (Jung, Park, Kim, Na, Kim & Kim, 2019; Venek, Scherer, Morency & Pestian, 2017). Sobre esto en particular, el estudio realizado por Venek, Scherer, Morency & Pestian (2017) buscó a través los datos del comportamiento verbal y no verbal durante la interacción paciente-terapeuta comparar e identificar rasgos comunes en adolescentes con tendencias y suicidas y no suicidas. Así algunos elementos relevantes identificados a través de la clasificación realizada por el algoritmo *Support Vector Machine* (SPV) son por ejemplo el hecho que la intervención verbal del clínico suele ser menor durante la entrevista, sin embargo, dicha entrevista tiende a ser más extensa en duración; además se evidenció que el terapeuta interrumpe significativamente menos a pacientes con rasgos o histórico de intentos suicidas, comparado con adolescentes no suicidas. Entretanto, los pacientes con tendencias suicidas repetían con mayor frecuencia el pronombre personal “yo” e hicieron referencia en menos ocasiones a emociones positivas. En tal sentido, aunque diversos estudios relacionados con riesgo o tentativas suicidas coinciden en resaltar la importancia de rasgos individuales como dificultad en el manejo de emociones, estrés, depresión o ansiedad existen también variables contextuales que deben ser analizadas como por ejemplo el hecho de pertenecer a algún grupo minoritario ya sea por orientación sexual, racial o económica. (Jakobsen, Larsen, & Horwood, 2017; Smith, Wang, CarterFox, & Hooley, 2020). Por ende los estudios analizados en la presente revisión sistemática profundizan principalmente en fenómenos individuales, pero parecen carecer de profundidad al analizar fenómenos contextuales y es en ese contexto donde los modelos de Inteligencia artificial y Aprendizaje automático pueden representar un avance en la investigación de procesos psicológicos, considerando su robustez para analizar complejas, diversas y amplias cantidades de variables y datos.

Por otro lado, aunque únicamente un estudio abordó el tema de violencia escolar con modelos de Aprendizaje automático, los autores de la presente revisión consideran importante identificarla como una categoría relevante, tomando en cuenta la importancia del fenómeno y lo incipiente de las producciones académicas que han utilizado dichos métodos de análisis hasta la fecha. Así el estudio realizado

por Barzman, et. al (2018) resalta la importancia de reducir la subjetividad en los modelos de predicción aumentando la validez predictiva en los estudios prospectivos de investigación. Aunque los autores resaltan la precisión de los datos analizados con los algoritmos de Aprendizaje automático utilizados, algunos de los resultados parecen ser contradictorios con lo observado predominantemente en la literatura sobre violencia escolar, al referir que la raza no está asociada con violencia escolar y que en ese sentido los ingresos económicos suelen explicar con más claridad el fenómeno. Otro elemento que llama la atención y parece contradecir de igual manera lo observado en estudios previos es el sexo y los actos de violencia escolar, considerando que la mayoría de los estudios apuntan a que el hecho de ser hombre es una variable comúnmente asociada con el hecho de ser víctima-victimario en los procesos de violencia escolar (Brewer, Thomas, & Higdon, 2018; Kim, Sanders, Makubuya & Yu, 2020).

Finalmente, aunque queda claro que la aplicación de abordajes de inteligencia artificial y Aprendizaje automático son insuficientes para explicar la complejidad de los fenómenos psicológicos, es también importante destacar que su utilización en las diversas ramas de las ciencias psicológicas abre las puertas para más eficaces formas de identificación de fenómenos, los cuales al ser identificados precozmente pueden contribuir a implementar acciones preventivas más precisas, oportunas y de gran escala.

## REFERENCIAS

- Ahulu, L. D., Gyasi-Gyamerah, A. A., & Anum, A. (2020). Predicting risk and protective factors of generalized anxiety disorder: A comparative study among adolescents in Ghana. *International Journal of Adolescence and Youth*, 25(1), 574–584. <https://doi.org/10.1080/02673843.2019.1698440>
- Barzman, D., Ni, Y., Griffey, M., Bachtel, A., Lin, K., Jackson, H., ... & DelBello, M. (2018). Automated risk assessment for school violence: a pilot study. *Psychiatric quarterly*, 89(4), 817-828.
- Breiman, L. (2001). Statistical modeling: The two cultures (with comments and a rejoinder by the author). *Statistical Science* 16, 199–231.
- Brewer, N. Q., Thomas, K. A., & Higdon, J. (2018). Intimate partner violence, health, sexuality, and academic performance among a national sample of undergraduates. *Journal of American College Health*, 66(150), 01–29.
- Correll, C.U., Galling, B., Pawar, A., Krivko, A., Bonetto, C., Ruggeri, M., ... & Kane, J.M. (2018). Comparison of early intervention services vs treatment as usual for early-phase psychosis: a systematic review, meta-analysis, and meta-regression. *JAMA Psychiatry*, 75, 555–565.

- Cummings, C. M., Caporino, N. E., & Kendall, P. C. (2014). Comorbidity of anxiety and depression in children and adolescents: 20 years after. *Psychological bulletin*, 140(3), 816.
- Dwyer, D., & Koutsouleris, N. (2022). Annual Research Review: Translational Machine Learning for child and adolescent psychiatry. *Journal of Child Psychology and Psychiatry*.
- Essau, C. A., Lewinsohn, P. M., Lim, J. X., Moon-ho, R. H., & Rohde, P. (2018). Incidence, recurrence and comorbidity of anxiety disorders in four major developmental stages. *Journal of affective disorders*, 228, 248-253.
- Franklin JC, Ribeiro JD, Fox KR, Bentley KH, Kleiman EM, Huang X, et al. (2017). Risk factors for suicidal thoughts and behaviors: A meta-analysis of 50 years of research. *Psychol Bull.* 2017;143(2):187–232. pmid:27841450.
- Gálvez-Nieto, J. L., Vera-Bachmann, D., Trizano-Hermosilla, Í., Polanco, K., & Salvo, S. (2018). Propiedades Psicométricas de la Versión Reducida de la Escala de Valores para el Desarrollo Positivo Adolescente (EVDPA-R) en Estudiantes Chilenos. *Revista Colombiana de Psicología*, 27(2), 69-84.
- Jakobsen, I. S., Larsen, K. J., & Horwood, J. L. (2017). Suicide risk assessment in adolescents – C-SSRS, K10, and READ. *Crisis: The Journal of Crisis Intervention and Suicide Prevention*, 38(4), 247–254. <https://doi.org/10.1027/0227-5910/a000450>
- Kim, J.-W., Sharma, V., & Ryan, N. D. (2015). Predicting methylphenidate response in ADHD using Machine learning approaches. *The International Journal of Neuropsychopharmacology*, 18(11), pyv052. <https://doi.org/10.1093/ijnp/pyv052>.
- Kim, Y. K., Sanders, J. E., Makubuya, T., & Yu, M. (2020). Risk factors of academic performance: Experiences of school violence, school safety concerns, and depression by gender. *Child & Youth Care Forum*. Advance online publication. <https://doi.org/10.1007/s10566-020-09552-7>
- Luca, M., Kleinberg, J. and Mullainathan, S. (2016). 'Algorithms need managers, too', *Harvard Business Review*, Vol. 104, pp. 96-101.
- Marín, O. (2016). Developmental timing and critical windows for the treatment of psychiatric disorders. *Nature Medicine*, 22, 1229-1238.
- Maulud, D., & Abdulazeez, A. M. (2020). A review on linear regression comprehensive in machine learning. *Journal of Applied Science and Technology Trends*, 1(4), 140-147.
- McCrory, C., Dooley, C., Layte, R., & Kenny, R. A. (2015). The lasting legacy of childhood adversity for disease risk in later life. *Health Psychology*, 34(7), 687.
- McLeod, G. F., Horwood, L. J., & Fergusson, D. M. (2016). Adolescent depression, adult mental health and psychosocial outcomes at 30 and 35 years. *Psychological medicine*, 46(7), 1401-1412.

- McLoughlin, A. B., Gould, M. S., & Malone, K. M. (2015). Global trends in teenage suicide: 2003–2014. *QJM: An International Journal of Medicine*, 108(10), 765-780.
- Montes, K. S., Witkiewitz, K., Pearson, M. R., & Leventhal, A. M. (2019). Tobacco, Alcohol and marijuana expectancies as predictors of substance use initiation in adolescence: A longitudinal examination. *Psychology of Addictive Behaviors*, 33(1), 26–34. <https://doi.org/10.1037/adb0000422>
- Morrow, A. S., Campos Vega, A. D., Zhao, X., & Liriano, M. M. (2020). Leveraging machine learning to identify predictors of receiving psychosocial treatment for Attention Deficit/Hyperactivity Disorder. *Administration and Policy in Mental Health and Mental Health Services Research*, 47(5), 680-692.
- Nock, M. K., Green, J. G., Hwang, I., McLaughlin, K. A., Sampson, N. A., Zaslavsky, A. M., & Kessler, R. C. (2013). Prevalence, correlates, and treatment of lifetime suicidal behavior among adolescents: Results from the National Comorbidity Survey Replication Adolescent Suppl. *Journal of the American Medical Association Psychiatry*, 70, 300–310. <https://doi.org/10.1001/2013.jamapsychiatry.55>
- Patel, V., Flisher, A. J., Hetrick, S., & McGorry, P. (2007). Mental health of young people: A global public-health challenge. *The Lancet*, 369, 1302–1313.
- Paus, T., Keshavan, M., & Giedd, J. N. (2008). Why do many psychiatric disorders emerge during adolescence? *Nature Reviews Neuroscience*, 9(12), 947-957.
- Pine, D. S., Cohen, P., Gurley, D., Brook, J., & Ma, Y. (1998). The risk for early-adulthood anxiety and depressive disorders in adolescents with anxiety and depressive disorders. *Archives of general psychiatry*, 55(1), 56-64.
- Rajula, H. S. R., Verlatto, G., Manchia, M., Antonucci, N., & Fanos, V. (2020). Comparison of conventional statistical methods with machine learning in medicine: diagnosis, drug development, and treatment. *Medicina*, 56(9), 455.
- Riley, E. N., Rukavina, M., & Smith, G. T. (2016). The reciprocal predictive relationship between high-risk personality and drinking: An 8-wave longitudinal study in early adolescents. *Journal of Abnormal Psychology*, 125(6), 798–804. <https://doi.org/10.1037/abn0000189>
- Sagi, E. (2019). Taming big data: Applying the experimental method to naturalistic data sets. *Behavior Research Methods*, 51(4), 1619-1635.
- Sansone, D. (2019). Beyond early warning indicators: High school dropout and machine learning. *Oxford Bulletin of Economics and Statistics*, 81(2), 456–485. <https://doi.org/10.1111/obes.12277>
- Smith, D. M., Wang, S. B., Carter, M. L., Fox, K. R., & Hooley, J. M. (2020). Longitudinal predictors of self-injurious thoughts and behaviors in sexual and gender minority adolescents. *Journal of Abnormal Psychology*, 129(1), 114–121. <http://dx.doi.org/10.1037/abn0000483>
- Volkaert, B., Wante, L., Van Beveren, M. *et al.* Training Adaptive Emotion Regulation Skills in Early Adolescents: The Effects of Distraction, Acceptance, Cognitive

Reappraisal, and Problem Solving. *Cogn Ther Res* 44, 678–696 (2020).  
<https://doi.org/10.1007/s10608-019-10073-4>

## ANEXO

Tabla 1. Principales características de los estudios seleccionados

N°	Título / Autor	Objetivos	Tipo de estudio y método	País Año	Participantes	Principales resultados	Mecanismo de validación
1	A Machine learning Approach to College Drinking Prediction and Risk Factor Identification Bi, J., Sun, J., Wu, Y., Tennen, H., & Armeli, S.	Analizar datos recolectados previamente y que pudieran ofrecer información para la identificación de patrones e hipótesis de problemas asociados al consumo de alcohol en jóvenes.	Estudio longitudinal que utilizó regresión logística y Support Vector Machine (SVM) como complemento para el análisis de los resultados.	EE.UU. 2013	530 estudiantes universitarios matriculados en el curso de introducción a la psicología en la Universidad de Connecticut que reportaron haber bebido alcohol al menos dos veces en el último mes.	El estudio resalta que abordajes de Aprendizaje automático ofrecen mejor capacidad de generalización y mayor capacidad predictiva al ser comparados con modelos tradicionales como el análisis de regresión logística.	Validación cruzada 3-k
2	Adolescent Suicidal Risk Assessment in Clinician-Patient Interaction Venek, V., Scherer, S., Morency, L. P., & Pestian, J.	A través de las características de la conversación, información verbal, no verbal y acústica, buscar segmentar a los participantes adolescentes en: no suicidas, suicidas y suicidas con más de un intento.	Estudio transversal. Utilizó inicialmente Análisis de Varianza y posteriormente el algoritmo de clasificación Support Vector Machine (SVM), para mejorar el desempeño de los datos se utilizó el algoritmo AdaBoost.	EE.UU. 2017	30 adolescentes que habían acudido al servicio de urgencias con tendencias suicidas y 30 sin tendencias suicidas.  De ellos, 13 habían tenido más de un intento y 17 con un solo intento.	Algunos de los resultados relevantes: Adolescentes suicidas suelen utilizar con mayor frecuencia el pronombre YO, las sesiones suelen durar más y las pausas o silencios suelen ser más prolongados respecto de los adolescentes no suicidas.	Lachenbruch's holdout validation
3	Automated Risk Assessment for School Violence: a Pilot Study Barzman, D., Ni, Y., Griffey, M., Bachtel, A., Lin, K., Jackson, H., ... & DelBello,	Por medio de entrevistas y las escalas de medición de riesgo de violencia escolar adaptadas previamente por los autores (Brief Rating of Aggression by Children and Adolescents BRACHA, School Safety Scale SSS y el Psychiatric Intake Response Center PIRC),	Estudio transversal que utilizó análisis estadísticos tradicionales como Chi cuadrado, correlación de Pearson. Con las entrevistas transcritas, se usó un algoritmo de regresión logística multivariante con	EE.UU. 2018	103 estudiantes de secundaria y preparatoria que provenían de 74 escuelas de Ohio, Kentucky, Indiana y Tennessee.	Tanto la escala BRACHA como la escala SSS identificaron riesgo de agresión hacia los otros, pero no de autoagresión entre los adolescentes. Con el contenido de entrevistas transcritas, el algoritmo de regresión logística multivariante con normalización L2 sugiere que podría ayudar a	Validación cruzada anidada 10-k

		identificar patrones del lenguaje y juicio asociados a la violencia escolar.	normalización L2, para entrenar el modelo.			minimizar la subjetividad clínica y maximizar la validez predictiva en la práctica clínica.  El algoritmo de Aprendizaje Automático fue preciso al medir el riesgo de violencia escolar en comparación con los juicios clínicos realizados por el psiquiatra forense.	
4	Dissociable psychosocial profiles of adolescent substance users.  Fitzgerald, A., Mac Giollabhui, N., Dolphin, L., Whelan, R., & Dooley, B.	Examinar el papel del individuo, la familia, la escuela, los compañeros y el entorno social sobre el uso de alcohol, tabaco y cannabis.	Estudio transversal (datos recolectados entre febrero y octubre de 2011). Se usaron medidas de autoinforme con un solo ítem y los datos fueron analizados usando regresión logística con la regularización Elastic Net.	Irlanda 2018	6,062 estudiantes de una muestra de las 732 escuelas secundarias en Irlanda.	Los perfiles de consumo de alcohol se destacan por la contribución de múltiples dominios (Individuales, familiares, escolares y sociales).  En contraste, el consumo de tabaco se caracterizó por un pequeño número de variables individuales, incluido el ser mujer y tener una baja auto percepción académica.  El consumo de cannabis mostró principalmente la contribución de factores de riesgo individuales, en particular ser hombre y experimentar sentimientos de ira.	Validación cruzada 10-k
5	Prediction models for high risk of suicide in Korean adolescents using machine learning techniques.  Jung, J. S., Park, S. J., Kim, E. Y., Na, K. S., Kim, Y. J., & Kim, K. G	Desarrollar un modelo de predicción con base en técnicas de Aprendizaje automático para identificar a adolescentes en alto riesgo de suicidio (jóvenes con ideación o tentativas suicidas).	El proceso de clasificación se realizó utilizando técnicas Aprendizaje automático como regresión logística (LR), Random Forest (RF), Support Vector Machine (SVM), extreme gradient boosting (XGB), Artificial Neural Network (ANN)	South Korea 2019	Se evaluó a 59,984 adolescentes por medio de un conjunto de datos tomados de la Encuesta nacional basada en la conducta de riesgo juvenil de Korea. (KYRBWS).	Teniendo en cuenta 26 variables como predictores, la precisión de los modelos de aprendizaje automático para predecir el suicidio fue comparable con la de Regresión logística. La precisión fue mejor en XGB (79,0 %), seguida de SVM (78,7 %), LR (77,9 %), RF (77,8 %) y ANN (77,5 %).	Validación cruzada 5-k