

Predicción del rendimiento académico en educación secundaria mediante el análisis de árboles de decisión

Academic achievement prediction in secondary education by decision tree analysis

Israel Villarrasa-Sapiña ¹ 

Xavier García-Massó ^{2*} 

Encarnación Liébana ³ 

Gonzalo Monfort Torres ⁴ 

¹ Universidad Internacional de La Rioja, Spain

² Universitat de València, Spain

³ Universitat Catòlica de València, Spain

⁴ Universitat Politècnica de València, Spain

* Autor de correspondencia. E-mail: xavier.garcia@uv.es

Cómo referenciar este artículo/ How to reference this article:

Villarrasa-Sapiña, I., García-Massó, X., Liébana, E., & Monfort Torres, G. (2024). Predicción del rendimiento académico en educación secundaria mediante el análisis de árboles de decisión [Academic achievement prediction in secondary education by decision tree analysis]. *Educación XX1*, 27(1), 253-279. <https://doi.org/10.5944/educxx1.33351>

Fecha de recepción: 08/03/2022

Fecha de aceptación: 17/04/2023

Publicado online: 02/01/2024

RESUMEN

El objetivo del presente estudio fue desarrollar un modelo de predicción del rendimiento académico (éxito o fracaso escolar) mediante la aplicación de un análisis de árbol de decisión. Se realizó un estudio transversal para diseñar un sistema de detección temprana del fracaso escolar. Participaron 219 adolescentes (de 14 a 16 años) y se recabó información

de su estatus socioeconómico, percentil de índice de masa corporal (IMC), actividad física, tiempo de ocio frente a pantallas, niveles de disfrute, esperanza, ira, ansiedad, aburrimiento, compromiso conductual, compromiso emocional, compromiso cognitivo, rendimiento escolar autopercebido e intención de ir a la universidad, como variables de entrada en el análisis del árbol de decisión. Se encontraron 6 grupos de fracaso y 3 de éxito capaces de predecir el rendimiento académico. Se obtuvo una buena precisión en los conjuntos de datos de entrenamiento (80.11 %) y validación (81.40 %) del árbol de decisión. Es posible predecir el fracaso o el éxito académico mediante la evaluación del estado de peso, la actividad física, la ira y la esperanza durante la asistencia a la escuela, la intención de ir a la universidad y el rendimiento escolar autopercebido.

Palabras clave: educación secundaria, rendimiento académico, predicción, nivel de actividad física, árbol de decisión

ABSTRACT

The aim of the present study was to develop a predictive model of academic achievement (school success or failure) by applying a decision tree analysis. A cross-sectional study was carried out to design a system for the early detection of academic failure. 219 adolescents (aged 14 to 16) participated and information on their socioeconomic status, body mass index (BMI) percentile, physical activity, leisure time spent in front of screens, enjoyment, hope, anger, anxiety, boredom, behavioral engagement, emotional engagement, cognitive engagement, self-perceived school performance and intention to go to university was collected as input variables in decision tree analysis. 6 failure and 3 success groups were found able to predict academic performance. Good accuracy was obtained in the training (80.11 %) and validation (81.40 %) datasets of the decision tree. It is possible to predict academic failure or success by assessing weight status, physical activity, anger and hope during school attendance, intention to go to university and self-perceived school performance.

Keywords: high schools, academic achievement, prediction, physical activity level, decision tree

INTRODUCCIÓN

El fracaso escolar es un término polisémico que suele asociarse a la no consecución de un objetivo académico, lo que suele significar no aprobar determinadas asignaturas o alcanzar una titulación mínima. Sin duda, poder prevenir estas situaciones reduciría la frustración de los alumnos y de las familias, y supondría un gran avance para la sociedad, de forma que todos los alumnos pudieran tener una mejor educación (Alexander et al., 1997, 2001; Cairns et al., 1989). Ante los fracasos académicos, varios autores han propuesto diferentes formas

de detección precoz basadas en diversos factores, como las emociones, la forma física, el sedentarismo o el compromiso académico, que se detallan a continuación (Alzina & Escoda, 2012; D'Mello et al., 2008; Pekrun et al., 2002; Weiner, 1982). Cabe señalar que estas formas de detección precoz se crean para detectar el fracaso escolar con la suficiente antelación para que los profesionales del sistema educativo puedan intervenir y mejorar la situación del alumno.

Un factor que ha demostrado ser muy relevante a la hora de predecir tanto el rendimiento académico como el fracaso escolar es el estatus socioeconómico de las familias de los alumnos (Parr & Bonitz, 2015; Trujillo-Torres et al., 2020). Concretamente, el fracaso escolar en España no se distribuye por igual en la estratificación socioeconómica, ya que la clase social del alumno puede afectar al fracaso y al rendimiento, puesto que el porcentaje de fracaso escolar es mayor en los niños de clase trabajadora que en los de clase media (Martínez-García, 2011).

Pekrun et al. (2002) definió las emociones del estudiante como la experiencia personal al realizar actividades académicas, identificándolas como una parte muy importante de su motivación personal para lograr el éxito académico y evitar el fracaso escolar. En la misma línea, D'Mello et al. (2008) destacaron que conocer las emociones de los alumnos era importante para llevar a cabo un buen proceso de enseñanza, debido a las relaciones entre cognición y emoción. Estas relaciones se han descrito explicando que los alumnos experimentan confusión cuando se enfrentan a obstáculos en sus objetivos, o detectan contradicciones, incongruencias o anomalías en el proceso de enseñanza (Festinger, 1962; Graesser & Olde, 2003). Si la confusión no se resuelve, puede desembocar en irritación, frustración, ira y, a veces, incluso rabia. Por eso es comprensible que múltiples estudios hayan llegado a la conclusión de que emociones negativas como la ira, la ansiedad y el aburrimiento se correlacionan negativamente con el rendimiento académico (Pekrun, 2006; Pekrun et al., 2011). Sin embargo, tampoco se ha comprobado que unos valores demasiado bajos sean positivos para emociones como la ira (Lane et al., 2005; Pekrun et al., 2011). Por otra parte, un alumno puede experimentar una serie de emociones positivas (como el disfrute) cuando se enfrenta a retos, descubre conocimientos y domina conceptos. Los estudiantes que participan activamente en el proceso de aprendizaje pueden tener una experiencia similar a la fluidez, cuando están tan absortos en el material que el tiempo y la fatiga desaparecen (Csikszentmihalyi, 2014). De hecho, se considera que una emoción positiva como la esperanza tiene potencial suficiente para reconducir a los alumnos con bajo rendimiento (Dixson, 2019).

Otro factor que se ha relacionado con el rendimiento académico es la práctica de actividad física (AF) y algunos conceptos relacionados como el sedentarismo o la condición física. En un estudio longitudinal, Pellicer-Chenoll et al. (2015) llegaron a la conclusión de que el grupo de estudiantes con mayor práctica de AF

y forma física presentaba un índice de masa corporal (IMC) más bajo y un mayor rendimiento académico en comparación con los compañeros que realizaban menos AF. A su vez, el clúster de estudiantes con menor AF mostraba niveles más bajos de forma física, mayor IMC y menor rendimiento académico que el resto de perfiles de estudiantes. Varios estudios han encontrado este tipo de relación entre la AF (Marques et al., 2017; Morales, Pellicer-Chenoll, et al., 2011; Rasberry et al., 2011; A. Singh et al., 2012; Sullivan et al., 2017) o condición física (Coe et al., 2013; Van Dusen et al., 2011; Wittberg et al., 2009) y rendimiento académico. Sin embargo, también hay estudios que sostienen que no existen pruebas concluyentes sobre los efectos beneficiosos de la AF en el rendimiento cognitivo y académico general de los estudiantes (Rasberry et al., 2011). Estas discrepancias entre los distintos estudios pueden deberse a las diferentes concepciones del rendimiento académico y a la forma de medir y considerar la práctica de AF. En general, la relación entre ambas variables se considera positiva o inexistente (Singh et al., 2019).

También se ha explorado la posibilidad de que los hábitos sedentarios influyan negativamente en el rendimiento académico (aparte de que la AF pueda influir positivamente). Peiró-Velert et al. (2014), observó la influencia del tiempo dedicado al uso sedentario de pantallas (por ejemplo: videojuegos, teléfonos móviles, televisión...) en el rendimiento académico. Los resultados mostraron que existía una relación inversamente proporcional entre el rendimiento académico y el uso de pantallas.

El compromiso de los estudiantes también se ha estudiado como posible factor que influye en el fracaso escolar. Carini et al. (2006), corroboró que el compromiso de los estudiantes está positivamente vinculado a resultados de aprendizaje deseables, como el pensamiento crítico y buenas calificaciones. Dogan (2015) analizó este factor dividiéndolo en tres dimensiones de compromiso: cognitivo, conductual y emocional. Sus resultados mostraron que el compromiso cognitivo predecía el rendimiento académico, pero el compromiso emocional y conductual no eran predictores. Otros estudios señalan que el compromiso conductual es importante para obtener resultados académicos positivos y evitar el abandono escolar (Connell & Wellborn, 1991; Finn, 1989).

Como se ha explicado, hoy en día se sabe que estos factores, de forma aislada, influyen en el rendimiento académico. Sin embargo, se ha publicado un número limitado de estudios con el objetivo de desarrollar un sistema de detección precoz del fracaso escolar utilizando la mayoría de los factores mencionados.

Casillas et al. (2012), examinó los efectos combinados de las variables predictoras en la estimación del fracaso escolar. Sus conclusiones destacan la importancia de utilizar varios factores predictores (es decir, psicosociales y conductuales) para lograr una estimación precisa de los alumnos en riesgo de abandono.

Davis et al. (2014), realizaron un estudio que evaluaba hasta qué punto varios aprendizajes de habilidades socioemocionales (motivación de autoeficacia académica, conexiones sociales, importancia de la escuela y gestión escolar, gestión de la angustia psicológica y emocional, y estrés académico) podían utilizarse como predictores de los resultados académicos. Sus resultados indicaron que la combinación de subescalas de aprendizaje socioemocional discriminaba eficazmente entre los estudiantes que progresaban positivamente hacia la graduación en la escuela secundaria y aquellos identificados como desertores escolares.

Zhang et al. (2018), centraron sus esfuerzos en averiguar el tipo de modelo predictivo que lograba la mejor precisión, realizando una comparación entre modelos de clasificación como el Naïve Bayes, las máquinas de vectores soporte, el árbol de decisión y el perceptrón multicapa, obteniendo mejores resultados con los dos últimos. De hecho, otros estudios han utilizado árboles de decisión como método de análisis para predecir el abandono escolar, y han encontrado resultados con una precisión relativamente alta en sus estimaciones (Quadri & Kalyankar, 2010; Veitch, 2004). Cabe destacar que este análisis para este tipo de estudio es muy relevante e importante en comparación con otros análisis predictivos debido a su eficacia y múltiples beneficios. Por ejemplo, el árbol de decisión puede manejar diferentes tipos de datos de entrada (es decir, nominales, numéricos y de texto), es fácil de entender y puede procesar valores erróneos del conjunto de datos, entre otros (Rokach & Maimon, 2014).

Como resumen de los factores más utilizados para predecir el fracaso o el éxito académico, nos remitimos a una de las revisiones más recientes, realizada por Alyahyan y Düştægör (2020), que proporciona información sobre las variables más utilizadas en este tipo de estudios, basándose en la definición de éxito académico de York et al. (2015). Según esta revisión, los factores demográficos y psicológicos de los alumnos han demostrado ser dos de los más utilizados para la predicción, junto con el rendimiento académico previo y los factores del entorno de los alumnos.

Como se puede observar, aunque la base teórica parece tener un fundamento sólido sobre los factores más influyentes, otros factores relevantes como la AF, el IMC o los hábitos sedentarios no han sido abordados, a pesar de que existe suficiente literatura que ha demostrado su influencia en el rendimiento académico. Esto podría deberse a que los estudios realizados con este tipo de variables se han centrado en examinar el rendimiento académico más que el fracaso o el aprobado escolar. Por este motivo, es muy posible que estos factores no hayan aparecido como relevantes porque no han sido analizados.

Como se ha explicado anteriormente, algunos estudios han utilizado métodos de análisis relativamente novedosos (como el árbol de decisión) para predecir el fracaso escolar utilizando combinaciones de variables psicológicas y demográficas. Sin embargo, no se han publicado estudios sobre el desarrollo de sistemas

de detección precoz del fracaso escolar utilizando combinaciones no sólo de características psicológicas y demográficas, sino también de variables relacionadas con los estilos de vida, como la práctica de AF y actividades sedentarias. Por ello, consideramos que nuestro trabajo es novedoso, ya que incluye los factores más influyentes en la predicción del éxito académico de los alumnos estudiados, de acuerdo con Alyahyan y Düşteğör (2020), más otros elementos y aspectos que hasta ahora no se habían considerado de forma interrelacionada, como el IMC, la AF o los hábitos sedentarios. Además, la aplicación de un análisis multifactorial y no lineal, como el árbol de decisión, evita las limitaciones que pueden conllevar los análisis lineales (por ejemplo, la reducción de la potencia estadística cuando se añaden muchos factores o la multicolinealidad) y, además, este análisis proporciona resultados de clasificación y predicción fácilmente interpretables de forma visual. Por último, la importancia de este tipo de estudios radica en que, si se encuentra una combinación de factores que puedan identificar el fracaso escolar, a su vez, se pueden promover acciones que ayuden a evitarlos y así buscar el éxito.

El objetivo de este estudio fue, por tanto, desarrollar un modelo predictivo del rendimiento académico (éxito o fracaso escolar) mediante el análisis de árboles de decisión, utilizando las emociones sobre la asistencia a la escuela, el compromiso escolar, la AF, el tiempo de ocio dedicado a actividades sedentarias que requieren el uso de pantallas, las características sociodemográficas y las variables de ajuste escolar.

MATERIAL Y MÉTODOS

Diseño del estudio y participantes

Se realizó un estudio transversal para diseñar un sistema de detección precoz del fracaso escolar en alumnos de tercero y cuarto de educación secundaria en España. Para ello, los estudiantes completaron una serie de cuestionarios al inicio de un trimestre para medir la AF, las horas dedicadas a actividades sedentarias frente a pantallas, el estatus socioeconómico, las emociones relacionadas con la asistencia a clase, el compromiso escolar y el ajuste escolar. Estas variables junto con el IMC se utilizaron como variables de entrada para diseñar un árbol de clasificación para predecir los resultados académicos (éxito o fracaso) al final del trimestre (variable de salida).

La muestra estaba compuesta por 219 adolescentes (de 14 a 16 años). Los participantes fueron reclutados en los centros de enseñanza secundaria obligatoria de Valencia (España). Los criterios de inclusión fueron: i) tener entre 14 y 16 años (ambos inclusive), ii) no estar incapacitados neurológica o intelectualmente para comprender y cumplimentar los cuestionarios, y iii) negativa de sus progenitores a participar en el estudio. Las características de los participantes se recogen en la Tabla 1.

Tabla 1

Características de los estudiantes

Variable	Muestra total (n=219)	Hombres (n=100)	Mujeres (n=119)
Edad (años)	14.94 (0.79)	15.05 (0.79)	14.84 (0.77)
Altura (cm)	167.95 (8.82)	173.36 (7.96)	163.42 (6.73)
Peso (kg)	61.52 (11.16)	66.52 (11.91)	57.26 (8.49)
Percentil IMC	61.93 (24.94)	63.77 (26.62)	60.38 (23.45)

Nota. Datos expresados en forma de media (desviación estándar). IMC = Índice de masa corporal.

Los procedimientos aplicados en este estudio fueron aprobados previamente por el Comité Ético de la Universidad de Valencia (Código: 1503291), cumpliendo además los requisitos establecidos en la Declaración de Helsinki (1975, revisada posteriormente en 2008). Asimismo, los progenitores de los participantes dieron su consentimiento informado por escrito antes de participar en el experimento.

Medidas

Todas las medidas se tomaron en el instituto de los participantes, en su aula habitual. Los investigadores explicaron cómo rellenar el conjunto de cuestionarios y resolvieron las dudas de los alumnos. El tiempo necesario para rellenar todos los cuestionarios fue de entre 50 y 60 minutos.

Los cuestionarios se administraron al principio del trimestre. Esto permitió establecer una relación predictiva entre las variables de entrada y el fracaso escolar. Las calificaciones académicas de cada alumno obtenidas al final del trimestre fueron anonimizadas. Su rendimiento académico se codificó como «suspenso» si el alumno no superaba alguna de las asignaturas (fracaso escolar) o «aprobado» si el alumno aprobaba todas las asignaturas del trimestre (éxito académico).

Family Affluence Scale (FAS) II

Se trata de un cuestionario para determinar el estatus socioeconómico (ESE) de las familias de adolescentes en países europeos y norteamericanos (Currie et al., 2008). Se formularon cuatro preguntas objetivas para cuantificar los vehículos, los viajes de vacaciones, el dormitorio personal y los ordenadores, con el fin de estimar

la riqueza familiar. Las respuestas se codificaron de 0 (número mínimo de vehículos, viajes...) a 3 (número máximo de vehículos, viajes...) en función del número de respuestas posibles de cada ítem. La puntuación final de la escala se calculó como el valor medio de las puntuaciones de los ítems, siendo 0 el ESE más bajo y 2.25 el más alto.

Percentil Índice Masa Corporal

Los participantes declararon su peso y estatura y se calculó el IMC (kg/m²). Cabe señalar que Sherry et al. (2007), informaron de buenos resultados de validación del peso y la altura autoinformados en adolescentes. También se utilizaron tablas de crecimiento (Kuczmarski et al., 2000) para calcular el percentil de IMC (ajustado por edad y sexo).

Physical Activity Questionnaire for Adolescents (PAQ-A)

El PAQ-A fue validado por primera vez por Kowalski et al. (1997), como una versión modificada del cuestionario de AF para niños mayores. Se halló una buena validez convergente de este cuestionario para medir el nivel general de AF de estudiantes de secundaria. Posteriormente, Martínez-Gómez et al. (2009), validaron la versión española del PAQ-A obteniendo relaciones moderadas con los datos del acelerómetro ($\rho = 0.34 - 0.39$). Este cuestionario es útil para medir los niveles de AF desde muy baja hasta muy intensa en los últimos 7 días y es apropiado para adolescentes entre 13 y 18 años. Consta de ocho preguntas que evalúan diferentes aspectos de la AF realizada por el adolescente en diferentes periodos del día. El PAQ-A es un cuestionario sencillo, fácil de cumplimentar y de manejar en el entorno escolar. El resultado global de la prueba es una puntuación de 1 a 5 puntos (1 el valor más bajo y 5 el valor más alto) para determinar el nivel de AF realizado por cada adolescente.

Adolescent Sedentary Activity Questionnaire (ASAQ)

El ASAQ se utiliza para medir el tiempo dedicado a una serie de conductas sedentarias fuera del horario escolar durante una semana normal (Hardy et al., 2007). En este cuestionario los participantes responden a preguntas sobre quince hábitos sedentarios, con detalle acerca del tiempo (horas y minutos) a cada uno de ellos cada semana. Para este estudio, sólo se utilizaron los siete ítems referidos a las actividades sedentarias de ocio que requerían el uso de una tecnología de pantalla.

La puntuación final del tiempo de ocio dedicado a actividades sedentarias frente a pantallas se computó como el tiempo total reportado en cada uno de ellos en minutos.

Emotional Scales Questionnaire Related to Class Attendance (AEQ)

El AEQ fue diseñado por primera vez por Pekrun et al. (2011), para medir las emociones experimentadas por los estudiantes en relación con la asistencia a clase. El cuestionario completo consta de 24 escalas que miden varias emociones que se organizan en tres secciones para evaluar las emociones relacionadas con la clase, las relacionadas con el aprendizaje y las relacionadas con los exámenes. Los ítems se responden en base a una escala Likert de 5 puntos, donde el valor más bajo (1) se refiere a un total desacuerdo con la afirmación, mientras que el valor más alto (5) corresponde a un completo acuerdo con lo que expresa el ítem. Para este estudio se seleccionaron los 42 ítems de las escalas de disfrute (8 ítems), enfado (7 ítems), ansiedad (11 ítems), aburrimiento (9 ítems) y esperanza (7 ítems), durante la asistencia a clase (emociones relacionadas con la clase) en la versión española. Esta versión del cuestionario fue validada por Rosas (2015), quien encontró buenos parámetros en cuanto a fiabilidad, validez estructural y de constructo. La puntuación final de cada subescala se calcula como el valor medio de sus ítems.

School Engagement Measure Questionnaire (SEM)

El cuestionario SEM consta de 19 ítems, con un formato de escala Likert con un rango de 5 puntos (Fredricks & McColskey, 2012). La versión española del cuestionario fue validada por Díaz et al., (2016), quienes determinaron que 16 ítems se agrupaban en tres subescalas de compromiso: compromiso conductual (4 ítems; por ejemplo, «presto atención en clase»), emocional (5 ítems; por ejemplo, «me interesa el trabajo en el colegio») y cognitivo (7 ítems; por ejemplo «cuando leo un libro, me hago preguntas para asegurarme de que entiendo de qué trata»). El compromiso cognitivo se refiere al nivel de participación en la vida escolar y al desarrollo de habilidades de razonamiento complejas (Doğan, 2014). El concepto de compromiso emocional incluye el interés del alumno por la escuela, entendido como las reacciones del alumno en el aula y el nivel de interés, aburrimiento, infelicidad, felicidad y ansiedad del alumno (Skinner et al., 1990). Por último, el compromiso conductual está relacionado con la participación en actividades académicas, sociales o extraescolares. La puntuación de cada dimensión se calcula como el valor medio de los ítems asignados a esa subescala.

Brief Multidimensional School Adjustment Scale

La *Brief Multidimensional School Adjustment Scale* evalúa el grado de integración del adolescente en el entorno escolar (Rubia et al., 2010). Consta de 10 ítems con formato escala Likert con un rango de 6 puntos, que se dividen en 3 dimensiones: i) problemas de adaptación al entorno escolar (ítems 6, 7, 8, 9 y 10), ii) autopercepción del rendimiento escolar (ítems 1, 2 y 5), y iii) intención de ir a la universidad (ítems 3 y 4). De todas las dimensiones, sólo la autopercepción del rendimiento escolar y la intención de ir a la universidad fueron utilizadas y computadas como la media de los ítems que las componen. Se mide un autoconcepto positivo e integrado como estudiante, así como las expectativas de continuar con estudios superiores (Rubia et al., 2010).

Análisis Estadístico

El análisis de los datos se realizó mediante el programa Matlab R2018a (Mathworks Inc., Natick, EE. UU.). En primer lugar, se diseñó un árbol de clasificación para obtener un modelo de predicción del rendimiento académico (por ejemplo: los alumnos suspenden algunas asignaturas) utilizando como variables de entrada el ESE, el percentil de IMC, la AF, el tiempo de ocio frente a pantallas, el disfrute, la esperanza, la ira, la ansiedad, el aburrimiento, el compromiso conductual, el compromiso emocional, el compromiso cognitivo, el rendimiento escolar autopercebido y la intención de ir a la universidad.

El árbol de clasificación se validó utilizando una submuestra del conjunto de datos. Esta técnica consiste en dividir el número total de casos disponibles (es decir, 219) en dos conjuntos de datos: de entrenamiento (es decir, el 80% de los casos; $n = 176$) y de validación (es decir, el 20% de los casos; $n = 43$). No se encontraron diferencias significativas entre los conjuntos de datos de entrenamiento y validación en las variables utilizadas en este estudio. A continuación, se utilizó el conjunto de datos de entrenamiento para obtener el árbol de decisión, mientras que el conjunto de datos de validación se utilizó para verificar su validez. El árbol de decisión o clasificación es un método que divide la muestra en dos subgrupos utilizando una variable explicativa independiente. Para ello se establece un punto de corte de la variable explicativa, que divide la muestra en dos subnodos en función del valor de los sujetos en la variable, es decir, los casos que están por encima del umbral de un grupo y los casos que están por debajo forman el otro grupo. Este proceso se repite para cada subgrupo, hasta que todos los casos se clasifican correctamente.

En este estudio se utilizó como criterio de división el algoritmo CART con la desviación, que es un algoritmo binario que divide cada grupo en dos subgrupos. Además, para evitar un posible sobreentrenamiento (evitando una baja validez

externa) se estableció la condición, durante el proceso de entrenamiento, de que cada nodo tuviera al menos 10 casos para reducir el número final de nodos y divisiones. De este modo, el árbol de clasificación divide a los adolescentes en función de variables discriminantes para clasificar a todos los participantes según hayan aprobado todas las asignaturas (es decir, aprobados) o no (es decir, suspensos).

Una vez obtenido el modelo, se aplicó al conjunto de datos de validación para obtener variables de rendimiento de la clasificación. Se calcularon la precisión de la clasificación y los valores pronósticos suspendidos y aprobados. Estas variables se calcularon como se describe en las ecuaciones (1, 2 y 3).

$$\text{Precisión (\%)} = \frac{\text{Aprobados reales} + \text{suspensidos reales}}{\text{Muestra Total}} \cdot 100 \quad \text{Ec. 1}$$

$$\text{Valor pronóstico suspendidos (\%)} = \frac{\text{Suspensidos reales}}{\text{Suspensidos reales} + \text{Falsos Suspensidos}} \cdot 100 \quad \text{Ec. 2}$$

$$\text{Valor pronóstico aprobados (\%)} = \frac{\text{Aprobados reales}}{\text{Aprobados reales} + \text{Falsos Aprobados}} \cdot 100 \quad \text{Ec. 3}$$

RESULTADOS

En la sección siguiente se ofrecen los resultados obtenidos con respecto a los datos descriptivos de la muestra, así como los del árbol de decisión. Los datos descriptivos figuran en la Tabla 2.

Tabla 2

Datos descriptivos de las variables de entrada del árbol de clasificación

Variable	Estudiantes Aprobados	Estudiantes Suspensidos	Todos los estudiantes
Estatus Socioeconómico	2.62 [2.53–2.72]	2.43 [2.35–2.52]	2.52 [2.45–2.58]
Percentil IMC	58.43 [53.31–63.55]	64.66 [60.30–69.03]	61.93 [58.61–65.25]

Variable	Estudiantes Aprobados	Estudiantes Suspendidos	Todos los estudiantes
Actividad Física	2.43 [2.28–2.57]	2.38 [2.25–2.50]	2.40 [2.30–2.49]
Tiempo de ocio frente a pantallas	315.40 [266.04–364.76]	367.73 [325.59–409.87]	344.79 [312.79–376.79]
Disfrute	2.65 [2.48–2.81]	2.31 [2.15–2.46]	2.46 [2.34–2.57]
Esperanza	3.36 [3.16–3.57]	2.84 [2.65–3.02]	3.07 [2.93–3.21]
Ira	2.25 [2.10–2.40]	2.52 [2.33–2.71]	2.40 [2.28–2.53]
Ansiedad	1.90 [1.72–2.07]	2.25 [2.07–2.43]	2.09 [1.97–2.22]
Aburrimiento	2.75 [2.53–2.97]	2.99 [2.78–3.20]	2.88 [2.73–3.04]
Compromiso conductual	3.98 [3.82–4.14]	3.49 [3.34–3.64]	3.70 [3.59–3.82]
Compromiso emocional	3.24 [3.04–3.44]	2.90 [2.74–3.07]	3.05 [2.92–3.18]
Compromiso cognitivo	2.43 [2.27–2.59]	2.31 [2.18–2.44]	2.36 [2.26–2.46]
Rendimiento escolar autopercebido	4.13 [3.93–4.33]	3.21 [3.02–3.40]	3.61 [3.46–3.76]
Intención de ir a la universidad	5.1 [4.78–5.42]	3.73 [3.37–4.09]	4.33 [4.07–4.59]

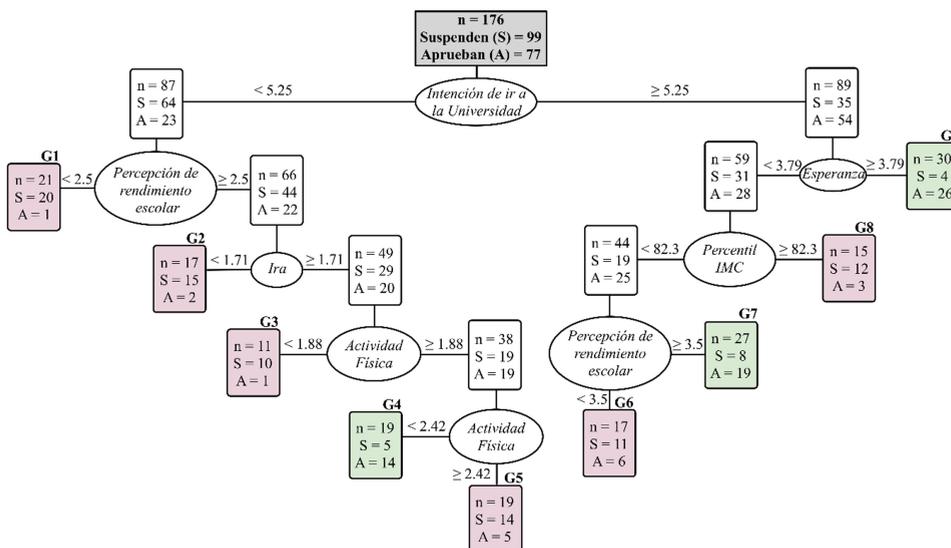
Nota. Datos expresados como media [95 % intervalo de confianza]. IMC = Índice Masa corporal.

El árbol de clasificación obtenido con el conjunto de datos de entrenamiento se muestra en la Figura 1. Los sujetos asignados en este conjunto de datos se dividen en subgrupos, utilizando las variables de entrada hasta llegar a los nodos terminales. Había 9 nodos terminales, cada uno de los cuales representaba un grupo

de participantes del conjunto de datos de entrenamiento. Estos nodos terminales se representan en la Figura 1 como G1, G2, [...], G9, y se utilizaron para clasificar a los estudiantes en función del fracaso académico (es decir, suspender) o del éxito (es decir, aprobar).

A modo de ejemplo, a continuación, se explica el primer nodo (es decir, G1) para ayudar a entender la Figura 1. G1 estaba formado por 21 alumnos, de los cuales 20 habían suspendido y 1 había aprobado la evaluación. Por tanto, se considera un nodo de fracaso escolar con una precisión de agrupación de participantes (a partir del conjunto de datos de entrenamiento) del 95 % (es decir, 20 dividido por 21). Dado que los 21 alumnos de este nodo presentaban una intención de ir a la universidad inferior a 5.25 puntos y un rendimiento escolar autopercebido inferior a 2.5 puntos, los futuros alumnos con estas características tendrán un alto riesgo de fracaso escolar que debería reducirse.

Figura 1
Árbol de clasificación para adolescentes en las fases de Aprobado (A) y Suspenso (S)



Nota. IMC = Índice de masa corporal.

La Tabla 3 muestra las características de cada nodo terminal para clarificar los resultados. En la primera columna, cada nodo terminal clasifica a los alumnos por éxito o fracaso académico, basándose en la proporción de personas que aprueban o suspenden del conjunto de datos de entrenamiento que se clasificaron en cada

uno de ellos. Las siguientes columnas muestran las variables de cada nodo terminal para describir las características de cada uno.

Tabla 3

Características de los nodos terminales del árbol de decisión

	IMCp	AF	Esperanza	Ira	REA	IU
G1 (Suspense)	–	–	–	–	<2.5	<5.25
G2 (Suspense)	–	–	–	<1.71	≥2.5	<5.25
G3 (Suspense)	–	<1.88	–	≥1.71	≥2.5	<5.25
G4 (Aprobado)	–	1.88–2.42	–	≥1.71	≥2.5	<5.25
G5 (Suspense)	–	≥2.42	–	≥1.71	≥2.5	<5.25
G6 (Suspense)	<82.3	–	<3.79	–	<3.5	≥5.25
G7 (Aprobado)	<82.3	–	<3.79	–	≥3.5	≥5.25
G8 (Suspense)	≥82.3	–	<3.79	–	–	≥5.25
G9 (Aprobado)	–	–	≥3.79	–	–	≥5.25

Nota. Percentil índice de masa corporal (IMCp), actividad física (AF), rendimiento escolar autopercebido (REA), intención de ir a la universidad (IU).

La Tabla 4 muestra el rendimiento del árbol de decisión para clasificar a los alumnos de los conjuntos de datos de entrenamiento y validación. De estos resultados se desprende que el árbol de decisión funcionó bien, tanto en los conjuntos de datos de entrenamiento (80.11% de precisión) como en los de validación (81.40% de precisión).

Tabla 4

Variables de rendimiento del árbol de decisión en los conjuntos de datos de entrenamiento y validación

	Árbol de decisión	
	Entrenamiento	Validación
Precisión (%)	80.11	81.40
VPA (%)	77.63	79
VPS (%)	82	83.33

Nota. VPA, Valor de pronóstico de aprobados; VPS, Valor de pronóstico de suspendidos.

DISCUSIÓN Y CONCLUSIONES

El objetivo de este manuscrito fue desarrollar un sistema de detección precoz del fracaso escolar mediante el análisis de árboles de decisión, utilizando como potenciales variables de entrada las emociones de los alumnos durante la asistencia a clase, el compromiso escolar, la intención de ir a la universidad, el rendimiento escolar autopercebido, el estatus socioeconómico, el percentil de IMC, la AF y el tiempo de ocio dedicado a actividades sedentarias que requieren el uso de pantallas (por ejemplo, ver la televisión). Como apoyan los resultados, el árbol de decisión demuestra una buena precisión de clasificación, ya que fue posible implementar este sistema en las escuelas como una forma fácil de detectar a los adolescentes en riesgo de suspender una asignatura. Este sistema sólo requiere que los estudiantes proporcionen el peso, la altura, el PAQ-A (8 ítems), los ítems relacionados con la ira (7 ítems) y la esperanza (7 ítems) durante la asistencia al colegio (es decir, del AEQ), así como los ítems relacionados con la intención de ir a la universidad (2 ítems) y el rendimiento escolar autopercebido (3 ítems) de la *Brief Multidimensional School Adjustment Scale*. Por tanto, aunque todas las variables de entrada utilizadas en este estudio son factores importantes en el rendimiento académico, no todas ellas son necesarias para predecir adecuadamente el fracaso escolar.

El fracaso escolar viene determinado por las dificultades que encuentran los estudiantes para alcanzar los objetivos educativos (Eisenberg et al., 2006; Enguita et al., 2010). Dado que en este influyen múltiples factores, es determinante conocer qué factores interfieren en mayor o menor medida en el fracaso escolar (Yu et al., 2018) para poder desarrollar sistemas de detección temprana. Hasta donde los autores conocen, varios estudios han realizado un análisis centrado en la predicción del fracaso escolar (Casillas et al., 2012; Respondek et al., 2017; Yu et al., 2018). Casillas et al. (2012) utilizan variables como el rendimiento académico, características psicosociales e indicadores conductuales, y Respondek et al. (2017) utilizan el control académico percibido y las emociones (por ejemplo, disfrute, aburrimiento y ansiedad). Sin embargo, la AF así como el uso de tecnología de medios de pantalla durante actividades de ocio sedentarias no se han tenido en cuenta, a pesar de la relación encontrada entre estas variables y el rendimiento académico (Morales, Pellicer-Chenoll, et al., 2011; Peiró-Velert et al., 2014).

Al incluir todas estas variables en un análisis, en nuestro estudio se obtuvo un árbol de decisión con una alta precisión (es decir, 80.11 % en los datos de entrenamiento y 81.4 % en los datos de validación). Estos resultados respaldan estudios anteriores, como el de Zhang et al. (2018), quienes observaron que los mejores métodos para predecir el rendimiento académico (cinco categorías) eran el árbol de decisión y el perceptrón multicapa (precisión del 57.41 % y el 62.04 % para el conjunto de datos de validación, respectivamente) frente a Naïve Bayes o las

máquinas de vectores de soporte (35,65 % y 48,61 %, respectivamente). Vairachilai (2020), concluyó que Naïve Bayes y el árbol de decisión tenían mayor precisión (77 y 71 %, respectivamente) que la máquina de vectores soporte (38 %). Por último, en el estudio de Ashraf et al. (2018), el árbol de decisión resultó ser el análisis predictivo más preciso (97.3 % de precisión). Cabe señalar que Ashraf et al., no aplicaron un procedimiento de validación cruzada y sus resultados (es decir, una precisión muy alta) podrían ser indicativos de un modelo sobreentrenado. En definitiva, puede decirse que el árbol de decisión es una herramienta adecuada para predecir el rendimiento académico y los parámetros de rendimiento de nuestro estudio son similares o superiores a los presentados en otros estudios con un propósito de clasificación similar.

Cabe señalar que el árbol de decisión obtenido en este estudio determina las características de los grupos de alumnos que aprobaron o suspendieron todas las asignaturas. Los resultados determinaron nueve grupos (nodos terminales del árbol de decisión), de los cuales seis representan grupos de alumnos que suspendieron alguna asignatura, mientras que los otros tres representan alumnos que aprobaron todas las asignaturas.

Así pues, existen seis combinaciones de características que conducen al fracaso escolar en los adolescentes. Por ejemplo, los alumnos asignados en G1 (figura 1) no presentaban una elevada intención de ir a la universidad y mostraban un bajo rendimiento escolar autopercebido, lo que se traducía en muchas probabilidades (en torno al 95 %) de suspender alguna de las asignaturas del curso (es decir, G1 en la figura 1). Pero lo realmente interesante es que, en la rama izquierda del árbol, los alumnos con intención de ir a la universidad inferior a 5.25 necesitan tener una puntuación relativamente alta en el rendimiento escolar autopercebido (≥ 2.5), experimentar al menos un nivel mínimo de enfado (≥ 1.88) durante algunos eventos presenciales, y realizar un nivel moderado de AF (de 1.88 a 2.42 puntos) para alcanzar el éxito académico (es decir, aprobar todas las asignaturas). Los alumnos con intención de ir a la universidad inferior a 5.25 que presenten cualquier otra combinación de las variables anteriores, tendrán una alta probabilidad de suspender alguna asignatura (es decir, G1, G2, G3 y G5 de la figura 1). Este hecho es muy relevante para sugerir estrategias que ayuden a estos alumnos, que tengan un deseo moderado de ir a la universidad, a superar las evaluaciones de bachillerato. Por ejemplo, si se detecta un alumno asignado en G3, podría ser interesante promover estilos de vida activos con el objetivo de aumentar moderadamente la AF realizada por el alumno con la esperanza de pasar de G3 (fracaso) a G4 (éxito).

Como se puede observar en este árbol de decisión, la intención de ir a la universidad es una variable de gran importancia, ya que es necesario conocer el resto de factores que influyen en los estudiantes. Por el contrario, en un estudio anterior no se había considerado como una variable predictora eficaz (Fernández-

Lasarte et al., 2019). Esto puede deberse a que en este estudio no se ha analizado este factor mediante análisis no lineales, que modulan el peso de cada variable predictora al interrelacionarla con el resto de las variables.

Estos resultados coinciden con los de estudios anteriores. Por un lado, puede parecer contradictorio que un valor muy bajo de ira sea una característica importante para determinar el fracaso escolar en algunos alumnos (es decir, G2), ya que múltiples estudios mencionan que un alto grado de ira es negativo para el rendimiento académico (Pekrun, 2006; Pekrun et al., 2011). No obstante, algunos estudios realizados, como los de Pekrun et al. (2011) y Lane et al. (2005) concluyeron que una cierta cantidad de ira puede tener efectos positivos en el rendimiento académico, mientras que tanto los valores muy altos como los muy bajos de ira son negativos para el rendimiento académico de los estudiantes. Esto podría explicarse por el control que los estudiantes tienen de su inteligencia emocional, ya que Parker et al. (2004) sugirieron que los individuos con un alto nivel de inteligencia emocional son conscientes de los efectos positivos de la ira sobre el rendimiento académico, y son capaces de regular su estado de ánimo para alcanzar estados adecuados y lograr el éxito académico. Por tanto, si se detecta un alumno con un nivel muy bajo de ira (es decir, asignado en G2), sería interesante analizar su habilidad en el dominio de la inteligencia emocional para que pueda controlar emociones como el aburrimiento (Pekrun et al., 2010), la desesperanza (Titz, 2001) y la ira, que, sin un control adecuado, pueden desestabilizar su estado emocional, y evitar así el fracaso escolar.

Por otro lado, los resultados extraídos de los Grupos 3, 4 y 5 corroboran que es necesario que los estudiantes tengan una vida moderadamente activa para tener éxito académico. El G3 (fracaso) reafirma que la falta de AF se correlaciona positivamente con el fracaso escolar (Pellicer-Chenoll et al., 2015) e, inversamente, el G4 (éxito) corrobora que la práctica moderada de AF se relaciona positivamente con el éxito académico (Morales, González, et al., 2011; Morales, Pellicer-Chenoll, et al., 2011; Pellicer-Chenoll et al., 2015). Estos resultados pueden explicarse por razones neurofisiológicas (Hillman et al., 2005; Tomporowski et al., 2008; van Praag, 2009), ya que el ejercicio físico mejora el flujo sanguíneo al cerebro y por tanto mejoran las funciones cognitivas, o por razones psicosociales (Sallis et al., 1999; Sigfúsdóttir et al., 2007), ya que la AF se asocia positivamente con la salud mental, la autoestima, el bienestar emocional y el autoconcepto, lo que puede influir positivamente en el rendimiento académico (Pellicer-Chenoll et al., 2015). Aun así, esta correlación entre AF y rendimiento académico no es del todo lineal, y Morales, Pellicer-Chenoll, et al. (2011) ya han sugerido que altos niveles de AF no conllevan mejoras en el rendimiento académico, lo que apoyaría los resultados encontrados en G5 (fracaso). En otras palabras, la relación entre AF y rendimiento escolar podría ser no lineal, ya que los niveles moderados son los más vinculados con los logros escolares (es decir, función cuadrática).

Considerando la rama derecha del árbol, cabe señalar que hay dos grupos de características que conducen al éxito académico: i) intención de ir a la universidad igual o superior a 5.25 y experimentar esperanza durante la asistencia a clase al menos hasta 3.79 puntos; y ii) intención de ir a la universidad igual o superior a 5.25, experimentar niveles bajos a moderados de esperanza durante la clase (es decir, <3.79), tener un percentil de IMC inferior a 82.3 y tener un alto rendimiento escolar autopercebido (es decir, ≥ 3.5). Una buena forma de ayudar a los estudiantes motivados para ir a la universidad es asegurarse de que experimentan esperanza durante la asistencia a clase (es decir, ayudar a estos estudiantes a ser asignados en G9) para mejorar drásticamente su éxito académico. Estos resultados están en consonancia con otros estudios, que observaron una correlación positiva entre la esperanza y el rendimiento académico y propusieron la aplicación de intervenciones para aumentar la esperanza académica debido a sus beneficios (Feldman y Kubota, 2015).

Sin embargo, si no se puede fomentar la esperanza a niveles altos en algunos estudiantes, se debe promover un estilo de vida saludable para controlar el percentil de IMC y evitar que los estudiantes tengan sobrepeso (es decir, $85 \leq$ percentil 94) u obesidad (>95 percentil) (Kuczmarski et al., 2000) para que no fracasen académicamente (es decir, G8). Estos resultados coinciden con estudios previos, en los que tener un mayor porcentaje de IMC puede provocar desconexión o falta de compromiso con el trabajo académico (Finn et al., 2018; Peterson et al., 2012). Aquellas personas con sobrepeso u obesidad pueden tener una menor productividad debido a problemas de salud, que pueden estar relacionados con cuestiones sociales, psicológicas y afectivas (Shaw et al., 2015). Dado que el sobrepeso y la obesidad infantil están relacionados con la AF, la dieta y el sedentarismo, aquellos estudiantes con estilos de vida saludables podrían, por tanto, tener un IMC más bajo y probablemente alcanzar un buen rendimiento académico (Pellicer-Chenoll et al., 2015). Por tanto, se corroboran los resultados obtenidos en estudios anteriores, confirmando que la obesidad es importante para determinar el rendimiento académico y, en consecuencia, que los estilos de vida saludables también son factores importantes. Sin embargo, aunque los hábitos saludables son importantes para el éxito académico, los resultados de este estudio determinan que el sistema educativo también debe intentar que los alumnos tengan un rendimiento escolar autopercebido de moderado a alto para conseguir el éxito académico (es decir, G7) y evitar el fracaso (es decir, G6), como se muestra en ambas ramas del árbol. Por lo tanto, si no existe una alta motivación para ir a la universidad y un alto sentimiento de esperanza, los profesores deberían fomentar más este aspecto promoviendo un rendimiento autopercebido, de moderado a alto, para evitar el fracaso escolar. Una forma de mejorar el rendimiento escolar autopercebido podría ser ajustar la dificultad de las tareas de clase al nivel de los alumnos, intentando encontrar tareas

que les resulten difíciles, pero al mismo tiempo accesibles y superables. De este modo se reduciría el estrés y se mantendría la salud psicológica y emocional (Solberg et al., 1998). Además, este factor es aún más relevante al observar los resultados de estudios previos, donde se ha encontrado que una disminución del rendimiento escolar autopercibido conlleva una disminución de las expectativas académicas (García-Escalera et al., 2020). Esto sugiere que, por ejemplo, si un alumno está en G6, si su rendimiento escolar autopercibido no asciende, podría pasar a G1, aumentando la probabilidad de fracaso escolar (véase la relación de suspensos y aprobados en estos grupos en la Figura 1 para entender este ejemplo). De hecho, Godoy et al., (2013) concluyen que una autopercepción académica negativa puede ser un factor determinante en el rendimiento académico, y mencionan la necesidad de trabajar la autopercepción de los estudiantes en la escuela porque esto podría promover la mejora del rendimiento académico (Moreira et al., 2016).

Como se ha demostrado en los párrafos anteriores, conocer las características de estos grupos no sólo sería interesante para comprender la naturaleza del fracaso en el proceso académico, sino también para implementar estrategias que ayuden a los estudiantes a tener éxito en la escuela. El árbol de decisión tiene así dos aplicaciones prácticas principales: detectar estudiantes en riesgo de fracaso académico y proporcionar una orientación individualizada para diseñar estrategias para evitar el fracaso.

Un punto fuerte de este estudio fue el uso de un enfoque de árbol de decisión para probar combinaciones multifactoriales que incluyen variables relacionadas con estilos de vida activos. Además, se llevó a cabo en una sola ciudad, por lo que sus hallazgos deben ser corroborados y extendidos a otros lugares con mayor número de participantes para aumentar la generalización de los resultados. Otra limitación es el número limitado de variables de entrada, ya que otras variables interesantes como los hábitos alimentarios o las relaciones sociales escolares podrían aumentar el rendimiento del árbol de decisión a la hora de clasificar a los estudiantes según sus resultados académicos.

Aplicación práctica

El modelo predictivo de rendimiento académico presentado en este estudio puede tener una transferencia práctica de conocimiento al sistema educativo, ya que podría usarse como herramienta para detectar el fracaso escolar. Esta herramienta facilitaría orientar la intervención de los profesionales de la educación para mejorar su situación (es decir, docentes y profesionales especialistas). Para implementarlo correctamente, la dirección del centro debería poner a disposición estos cuestionarios, así como una explicación de su uso. Además, para mejorar aún más su usabilidad, los responsables educativos podrían crear un curso de aprendizaje o

crear una aplicación web que explique, paso a paso, el procedimiento y proporcione retroalimentación para abordar cada situación detectada de posible fracaso académico. En cualquier caso, a continuación, se presenta una explicación detallada de cómo aplicar este sistema con las herramientas disponibles actualmente. Estos pasos sólo deben seguirse para aquellos estudiantes cuyas calificaciones sean bajas o que el docente/profesional considere que pueden necesitarlo:

1. Responda las preguntas 3 y 4 de la *Brief Multidimensional School Adjustment Scale*. Si el resultado es igual o superior a 5.25 acuda al punto 2.a. y si es inferior a 5.25 acuda al punto 2.b.
- 2.a. Responda los 7 ítems de esperanza del *Emotional Scales Questionnaire Related to Class Attendance* (AEQ). Si el resultado es inferior a 3.79, el profesor pesaría y mediría al estudiante para calcular el percentil del IMC. Si el percentil de IMC es igual o superior a 82.3, este estudiante necesita ser ayudado en base a estas características. En cambio, si el estudiante tiene un percentil inferior de 82.3, deberá responder las preguntas 1, 2 y 5 de la *Brief Multidimensional School Adjustment Scale*. Si el resultado es inferior a 3.5, este estudiante también necesita ayuda. De esta manera, con un máximo de 12 preguntas sencillas y el cálculo del percentil del IMC se podría determinar el rendimiento académico del estudiante y los motivos de su desempeño.
- 2.b. Responda las preguntas 1, 2 y 5 de la *Brief Multidimensional School Adjustment Scale*. Si el resultado es inferior a 2.5, este alumno necesita ayuda en función de estas características. Sin embargo, si el resultado es igual o superior a 2.5, este estudiante deberá responder los 7 ítems de ira del *Emotional Scales Questionnaire Related to Class Attendance* (AEQ). Si el resultado es inferior a 1.71, este estudiante necesita ser ayudado en base a estas características. Por último, si el resultado es igual o superior a 1.71, este estudiante deberá responder el *Physical Activity Questionnaire for Adolescents* (PAQ-A) (8 ítems). Si el resultado es inferior a 1.88 o igual o superior a 2.42, este estudiante necesita ser ayudado en base a estas características. De esta manera, con un máximo de 18 preguntas sencillas se podría determinar el rendimiento académico del estudiante y los motivos de su desempeño.

REFERENCIAS BIBLIOGRÁFICAS

- Alexander, K. L., Entwisle, D. R., & Horsey, C. S. (1997). From first grade forward: early foundations of high school dropout. *Sociology of Education*, 70(2), 87–107. <https://doi.org/10.2307/2673158>

- Alexander, K. L., Entwisle, D. R., & Kabbani, N. S. (2001). The dropout process in life course perspective: early risk factors at home and school. *Teachers College Record*, 103(5), 760–822. <https://doi.org/10.1111/0161-4681.00134>
- Alyahyan, E., & Düşteğör, D. (2020). Predicting academic success in higher education: literature review and best practices. *International Journal of Educational Technology in Higher Education*, 17(1), 3. <https://doi.org/10.1186/s41239-020-0177-7>
- Alzina, R. B., & Escoda, N. P. (2012). Educación emocional: estrategias para su puesta en práctica. *Avances en Supervisión Educativa*, 16, 16. <https://avances.adide.org>
- Ashraf, A., Anwer, S., & Khan, M. G. (2018). A Comparative study of predicting student's performance by use of data mining techniques. *American Scientific Research Journal for Engineering, Technology, and Sciences*, 44(1), 1.
- Cairns, R. B., Cairns, B. D., & Neckerman, H. J. (1989). Early school dropout: configurations and determinants. *Child Development*, 60(6), 1437–1452. <https://doi.org/10.2307/1130933>
- Carini, R. M., Kuh, G. D., & Klein, S. P. (2006). Student engagement and student learning: testing the linkages. *Research in Higher Education*, 47(1), 1–32. <https://doi.org/10.1007/s11162-005-8150-9>
- Casillas, A., Robbins, S., Allen, J., Kuo, Y.-L., Hanson, M. A., & Schmeiser, C. (2012). Predicting early academic failure in high school from prior academic achievement, psychosocial characteristics, and behavior. *Journal of Educational Psychology*, 104(2), 407–420. <https://doi.org/10.1037/a0027180>
- Coe, D. P., Peterson, T., Blair, C., Schutten, M. C., & Peddie, H. (2013). Physical fitness, academic achievement, and socioeconomic status in school-aged youth. *The Journal of School Health*, 83(7), 500–507. <https://doi.org/10.1111/josh.12058>
- Connell, J. P., & Wellborn, J. G. (1991). Competence, autonomy, and relatedness: a motivational analysis of self-system processes. In M. R. Gunnar & L. A. Sroufe (Eds.) *Self processes and development* (pp. 43–77). Lawrence Erlbaum Associates, Inc.
- Csikszentmihalyi, M. (2014). Toward a psychology of optimal experience. En M. Csikszentmihalyi (Ed.), *Flow and the foundations of positive psychology: the collected works of mihaly csikszentmihalyi* (pp. 209–226). Springer Netherlands. https://doi.org/10.1007/978-94-017-9088-8_14
- Currie, C., Molcho, M., Boyce, W., Holstein, B., Torsheim, T., & Richter, M. (2008). Researching health inequalities in adolescents: the development of the Health Behaviour in School-Aged Children (HBSC) family affluence scale. *Social Science & Medicine*, 66(6), 1429–1436. <https://doi.org/10.1016/j.socscimed.2007.11.024>
- Davis, A., Solberg, V. S., Baca, C. de, & Gore, T. H. (2014). Use of social emotional learning skills to predict future academic success and progress toward

- graduation. *Journal of Education for Students Placed at Risk*, 19(3–4), 169–182. <https://doi.org/10.1080/10824669.2014.972506>
- Díaz, E. R., Fernández, A. R., & Revuelta, L. R. (2016). Validation of the Spanish version of the School Engagement Measure (SEM). *The Spanish Journal of Psychology*, 19, 85–89.
- Dixon, D. D. (2019). Hope into action: how clusters of hope relate to success-oriented behavior in school. *Psychology in the Schools*, 56(9), 1493–1511. <https://doi.org/10.1002/pits.22299>
- D’Mello, S., Jackson, G., Craig, S., Morgan, B., Chip-Man, P., White, H., Person, N., Kort, B., Kaliouby, R., Picard, R., & Graesser, A. (2008). *AutoTutor detects and responds to learners affective and cognitive states*. Workshop on emotional and cognitive issues at the International Conference Intelligent Tutoring Systems. Montreal, Canada.
- Doğan, U. (2014). Validity and reliability of Student Engagement Scale. *Bartın University Journal of Faculty of Education*, 3(2), 390-403. <https://doi.org/10.14686/BUFEAD.201428190>
- Dogan, U. (2015). Student engagement, academic self-efficacy, and academic motivation as predictors of academic performance. *The Anthropologist*, 20(3), 553–561. <https://doi.org/10.1080/09720073.2015.11891759>
- Eisenberg, N., Damon, W., & Lerner, R. M. (2006). *Handbook of child psychology: social, emotional, and personality development (6th ed)*. John Wiley & Sons, Inc.
- Enguita, M. F., Martínez, L. M., & Gómez, J. R. (2010). *Fracaso y abandono escolar en España*. Fundación la Caixa.
- Feldman, D. B., & Kubota, M. (2015). Hope, self-efficacy, optimism, and academic achievement: distinguishing constructs and levels of specificity in predicting college grade-point average. *Learning and Individual Differences*, 37, 210–216. <https://doi.org/10.1016/j.lindif.2014.11.022>
- Fernandez-Lasarte, O., Goñi, E., Camino, I., & Zubeldia, M. (2019). Ajuste escolar y autoconcepto académico en la Educación Secundaria. *Revista de Investigación Educativa*, 37(1), 1. <https://doi.org/10.6018/rie.37.1.308651>
- Festinger, L. (1962). *A theory of Cognitive Dissonance*. Stanford University Press.
- Finn, J. D. (1989). Withdrawing from school. *Review of Educational Research*, 59(2), 117–142. <https://doi.org/10.3102/00346543059002117>
- Finn, K. E., Faith, M. S., & Seo, Y. S. (2018). School engagement in relation to body mass index and school achievement in a high-school age sample. *Journal of Obesity*, 2018, 3729318. <https://doi.org/10.1155/2018/3729318>
- Fredricks, J. A., & McColskey, W. (2012). The measurement of student engagement: a comparative analysis of various methods and student self-report instruments. En S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on*

- student engagement* (pp. 763–782). Springer US. https://doi.org/10.1007/978-1-4614-2018-7_37
- García-Escalera, J., Valiente, R. M., Sandín, B., Ehrenreich-May, J., & Chorot, P. (2020). Los efectos de un programa de prevención de la ansiedad y la depresión para adolescentes en variables educativas y de bienestar. *Revista de Psicodidáctica*, 25(2), 143–149. <https://doi.org/10.1016/j.psicod.2020.05.001>
- Godoy, J. A., Abrahão, R. C., & Halpern, R. (2013). Autopercepção de dificuldades escolares em alunos do ensino fundamental e médio em município do Rio Grande do Sul. *Aletheia*, 41, 121–133.
- Graesser, A. C., & Olde, B. A. (2003). How does one know whether a person understands a device? The quality of the questions the person asks when the device breaks down. *Journal of Educational Psychology*, 95(3), 524–536. <https://doi.org/10.1037/0022-0663.95.3.524>
- Hardy, L. L., Booth, M. L., & Okely, A. D. (2007). The reliability of the Adolescent Sedentary Activity Questionnaire (ASAQ). *Preventive Medicine*, 45(1), 71–74. <https://doi.org/10.1016/j.ypmed.2007.03.014>
- Hillman, C. H., Castelli, D. M., & Buck, S. M. (2005). Aerobic fitness and neurocognitive function in healthy preadolescent children. *Medicine and Science in Sports and Exercise*, 37(11), 1967–1974. <https://doi.org/10.1249/01.mss.0000176680.79702.ce>
- Kowalski, K. C., Crocker, P. R. E., & Kowalski, N. P. (1997). Convergent validity of the Physical Activity Questionnaire for Adolescents. *Pediatric Exercise Science*, 9(4), 342–352. <https://doi.org/10.1123/pes.9.4.342>
- Kuzmarski, R., Ogden, C., Grummer-Strawn, L., Flegal, K., Shumei, G., Wei, R., Mei, Z., Curtin, L., Roche, A., & Johnson, C. (2000). CDC growth charts: United States. *National Center for Health Statistics*, 314, 28.
- Lane, A. M., Whyte, G. P., Terry, P. C., & Nevill, A. M. (2005). Mood, self-set goals and examination performance: the moderating effect of depressed mood. *Personality and Individual Differences*, 39(1), 143–153. <https://doi.org/10.1016/j.paid.2004.12.015>
- Marques, A., Corrales, F. R. G., Martins, J., Catunda, R., & Sarmento, H. (2017). Association between physical education, school-based physical activity, and academic performance: a systematic review. *Retos: Nuevas Tendencias En Educación Física, Deporte y Recreación*, 31, 316–320.
- Martínez García, J. S. (2011). Género y origen social: diferencias grandes en fracaso escolar administrativo y bajas en rendimiento educativo. *Revista de Sociología de la Educación-RASE*, 4(3), 270–282.
- Martínez-Gómez, D., Martínez-de-Haro, V., Pozo, T., Welk, G. J., Villagra, A., Calle, M. E., Marcos, A., & Veiga, O. L. (2009). Fiabilidad y validez del cuestionario

- de actividad física PAQ-A en adolescentes españoles. *Revista Española de Salud Pública*, 83(3), 427–439.
- Morales, J., González, L.-M., Guerra, M., Virgili, C., & Unnithan, V. (2011). Physical activity, perceptual-motor performance, and academic learning in 9-to-16-years-old school children. *International Journal of Sport Psychology*, 42(4), 401–415.
- Morales, J., Pellicer-Chenoll, M., García-Masso, X., Gomis, M., & González, L.-M. (2011). Relation between physical activity and academic performance in 3rd-year secondary education students. *Perceptual and Motor Skills*, 113(2), 539–546.
- Moreira, B. B. G., Martins-Reis, V. de O., & Santos, J. N. (2016). Autopercepção das dificuldades de aprendizagem de estudantes do ensino fundamental. *Audiology — Communication Research*, 21. <https://doi.org/10.1590/2317-6431-2015-1632>
- Parker, J. D. A., Summerfeldt, L. J., Hogan, M. J., & Majeski, S. A. (2004). Emotional intelligence and academic success: examining the transition from high school to university. *Personality and Individual Differences*, 36(1), 163–172. [https://doi.org/10.1016/S0191-8869\(03\)00076-X](https://doi.org/10.1016/S0191-8869(03)00076-X)
- Parr, A. K., & Bonitz, V. S. (2015). Role of family background, student behaviors, and school-related beliefs in predicting high school dropout. *The Journal of Educational Research*, 108(6), 504–514. <https://doi.org/10.1080/00220671.2014.917256>
- Peiró-Velert, C., Valencia-Peris, A., González, L. M., García-Massó, X., Serra-Añó, P., & Devís-Devís, J. (2014). Screen media usage, sleep time and academic performance in adolescents: clustering a self-organizing maps analysis. *Plos One*, 9(6), e99478. <https://doi.org/10.1371/journal.pone.0099478>
- Pekrun, R. (2006). The control-value theory of achievement emotions: assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review*, 18(4), 315–341. <https://doi.org/10.1007/s10648-006-9029-9>
- Pekrun, R., Goetz, T., Daniels, L. M., Stupnisky, R. H., & Perry, R. P. (2010). Boredom in achievement settings: exploring control–value antecedents and performance outcomes of a neglected emotion. *Journal of Educational Psychology*, 102(3), 531–549. <https://doi.org/10.1037/a0019243>
- Pekrun, R., Goetz, T., Frenzel, A. C., Barchfeld, P., & Perry, R. P. (2011). Measuring emotions in students' learning and performance: the Achievement Emotions Questionnaire (AEQ). *Contemporary Educational Psychology*, 36(1), 36–48. <https://doi.org/10.1016/j.cedpsych.2010.10.002>
- Pekrun, R., Goetz, T., Titz, W., & Perry, R. (2002). Academic emotions in students' self-regulated learning and achievement: a program of qualitative and quantitative

- research. *Educational Psychologist*, 37, 91–105. https://doi.org/10.1207/S15326985EP3702_4
- Pellicer-Chenoll, M., Garcia-Massó, X., Morales, J., Serra-Añó, P., Solana-Tramunt, M., González, L.-M., & Toca-Herrera, J.-L. (2015). Physical activity, physical fitness and academic achievement in adolescents: a self-organizing maps approach. *Health Education Research*, 30(3), 436–448. <https://doi.org/10.1093/her/cyv016>
- Peterson, J. L., Puhl, R. M., & Luedicke, J. (2012). An experimental assessment of physical educators' expectations and attitudes: the importance of student weight and gender. *Journal of School Health*, 82(9), 432–440. <https://doi.org/10.1111/j.1746-1561.2012.00719.x>
- Quadri, M. M. N., & Kalyankar, D. N. V. (2010). Drop out feature of student data for academic performance using decision tree techniques. *Global Journal of Computer Science and Technology*, 2(10). https://globaljournals.org/GJCST_Volume10/gjcst_vol10_issue2_ver1_paper7.pdf
- Raspberry, C. N., Lee, S. M., Robin, L., Laris, B. A., Russell, L. A., Coyle, K. K., & Nihiser, A. J. (2011). The association between school-based physical activity, including physical education, and academic performance: a systematic review of the literature. *Preventive Medicine*, 52, S10–S20. <https://doi.org/10.1016/j.ypmed.2011.01.027>
- Respondek, L., Seufert, T., Stupnisky, R., & Nett, U. E. (2017). Perceived academic control and academic emotions predict undergraduate university student success: examining effects on dropout intention and achievement. *Frontiers in Psychology*, 8. <https://doi.org/10.3389/fpsyg.2017.00243>
- Rokach, L., & Maimon, O. (2014). *Data mining with decision trees: theory and applications (2ª ed.)*. World Scientific Publishing Co., Inc.
- Rosas, J. S. (2015). The Achievement Emotions Questionnaire-Argentine (AEQ-AR): internal and external validity, reliability, gender differences and norm-referenced interpretation of test scores. *Revista Evaluar*, 15(1), 1. <https://doi.org/10.35670/1667-4545.v15.n1.14908>
- Rubia, J. C. M. de la, Sosa, J. C. S., & González, M. E. V. (2010). Desarrollo de una escala multidimensional breve de ajuste escolar. *Revista Electrónica de Metodología Aplicada*, 15(1), 1. <https://doi.org/10.17811/rema.15.1.2010.1-11>
- Sallis, J. F., McKenzie, T. L., Kolody, B., Lewis, M., Marshall, S., & Rosengard, P. (1999). Effects of health-related physical education on academic achievement: Project SPARK. *Research Quarterly for Exercise and Sport*, 70(2), 127–134. <https://doi.org/10.1080/02701367.1999.10608030>
- Shaw, S. R., Gomes, P., Polotskaia, A., & Jankowska, A. M. (2015). The relationship between student health and academic performance: implications for school

- psychologists. *School Psychology International*, 36(2), 115–134. <https://doi.org/10.1177/0143034314565425>
- Sherry, B., Jefferds, M. E., & Grummer-Strawn, L. M. (2007). Accuracy of adolescent self-report of height and weight in assessing overweight status: a literature review. *Archives of Pediatrics & Adolescent Medicine*, 161(12), 1154–1161. <https://doi.org/10.1001/archpedi.161.12.1154>
- Sigfúsdóttir, I. D., Kristjánsson, A. L., & Allegrante, J. P. (2007). Health behaviour and academic achievement in Icelandic school children. *Health Education Research*, 22(1), 70–80. <https://doi.org/10.1093/her/cyl044>
- Singh, A. S., Saliasi, E., Berg, V. van den, Uijtdewilligen, L., Groot, R. H. M. de, Jolles, J., Andersen, L. B., Bailey, R., Chang, Y.-K., Diamond, A., Ericsson, I., Etnier, J. L., Fedewa, A. L., Hillman, C. H., McMorris, T., Pesce, C., Pühse, U., Tomporowski, P. D., & Chinapaw, M. J. M. (2019). Effects of physical activity interventions on cognitive and academic performance in children and adolescents: a novel combination of a systematic review and recommendations from an expert panel. *British Journal of Sports Medicine*, 53(10), 640–647. <https://doi.org/10.1136/bjsports-2017-098136>
- Singh, A., Uijtdewilligen, L., Twisk, J. W. R., van Mechelen, W., & Chinapaw, M. J. M. (2012). Physical activity and performance at school: a systematic review of the literature including a methodological quality assessment. *Archives of Pediatrics & Adolescent Medicine*, 166(1), 49–55. <https://doi.org/10.1001/archpediatrics.2011.716>
- Skinner, E. A., Wellborn, J. G., & Connell, J. P. (1990). What it takes to do well in school and whether I've got it: a process model of perceived control and children's engagement and achievement in school. *Journal of Educational Psychology*, 82(1), 22–32. <https://doi.org/10.1037/0022-0663.82.1.22>
- Solberg, V. S., Gusavac, N., Hamann, T., Felch, J., Johnson, J., Lamborn, S., & Torres, J. (1998). The Adaptive Success Identity Plan (ASIP): a career intervention for college students. *The Career Development Quarterly*, 47(1), 48–95. <https://doi.org/10.1002/j.2161-0045.1998.tb00728.x>
- Sullivan, R. A., Kuzel, A. H., Vaandering, M. E., & Chen, W. (2017). The association of physical activity and academic behavior: a systematic review. *Journal of School Health*, 87(5), 388–398. <https://doi.org/10.1111/josh.12502>
- Titz, W. (2001). *Emotionen von Studierenden in Lernsituationen: explorative analysen und entwicklung von selbstberichtskalen — anhang*. <https://epub.uni-regensburg.de/9862/>
- Tomporowski, P. D., Davis, C. L., Miller, P. H., & Naglieri, J. A. (2008). Exercise and children's intelligence, cognition, and academic achievement. *Educational Psychology Review*, 20(2), 111–131. <https://doi.org/10.1007/s10648-007-9057-0>

- Trujillo-Torres, J.-M., Hossein-Mohand, H., Gómez-García, M., Hossein-Mohand, H., & Hinojo-Lucena, F.-J. (2020). Estimating the academic performance of secondary education mathematics students: a gain lift predictive model. *Mathematics*, 8(12), 12. <https://doi.org/10.3390/math8122101>
- Vairachilai S, V. (2020). Student's academic performance prediction using machine learning approach. *International Journal of Advanced Science and Technology*, 29, 9s.
- Van Dusen, D. P., Kelder, S. H., Kohl, H. W., Ranjit, N., & Perry, C. L. (2011). Associations of physical fitness and academic performance among schoolchildren. *The Journal of School Health*, 81(12), 733–740. <https://doi.org/10.1111/j.1746-1561.2011.00652.x>
- van Praag, H. (2009). Exercise and the brain: something to chew on. *Trends in Neurosciences*, 32(5), 283–290. <https://doi.org/10.1016/j.tins.2008.12.007>
- Veitch, W. R. (11-14 de abril de 2004). *Identifying characteristics of high school dropouts: data mining with a decision tree model* [Ponencia de Congreso]. Annual Meeting of the American Educational Research Association, San Diego. <https://eric.ed.gov/?id=ED490086>
- Weiner, B. (1982). An attribution theory of motivation and emotion. *Series in Clinical & Community Psychology: Achievement, Stress, & Anxiety*, 223–245.
- Wittberg, R. A., Northrup, K. L., & Cottrel, L. (2009). Children's physical fitness and academic performance. *American Journal of Health Education*, 40(1), 30–36.
- York, T. T., Gibson, C., & Rankin, S. (2015). Defining and measuring academic success. *Practical Assessment, Research & Evaluation*, 20(5). <https://eric.ed.gov/?id=EJ1059739>
- Yu, L. C., Lee, C. W., Pan, H. I., Chou, C. Y., Chao, P. Y., Chen, Z. H., Tseng, S. F., Chan, C. L., & Lai, K. R. (2018). Improving early prediction of academic failure using sentiment analysis on self-evaluated comments. *Journal of Computer Assisted Learning*, 34(4), 358–365. <https://doi.org/10.1111/jcal.12247>
- Zhang, X., Xue, R., Liu, B., Lu, W., & Zhang, Y. (28-30 de julio de 2018). *Grade prediction of student academic performance with multiple classification models*. [Ponencia de Congreso]. 14th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), China. <https://doi.org/10.1109/FSKD.2018.8687286>

