Predicción del rendimiento académico estudiantil con redes neuronales artificiales

- **Víctor D. Gil-Vera**^{1,2*} **y Catalina Quintero-López**³
 (1) Universidad Católica Luis Amigó, Facultad de Ingeniería de Sistemas, Medellín Colombia
- (2) Universidad Católica Luis Amigó, Grupo de Investigación SISCO, Medellín Colombia

(Correo-e: victor.gilve@amigo.edu.co)

- (3) Universidad Católica Luis Amigó, Grupo de Investigación en Neurociencias Básicas y Aplicadas, Medellín Colombia (Correo-e: catalina.quinterolo@amigo.edu.co)
- * Autor a quien debe ser dirigida la correspondencia

Recibido Jun. 10, 2021; Aceptado Ago. 9, 2021; Versión final Ago. 26, 2021, Publicado Dic. 2021

Resumen

Este trabajo presenta una red neuronal artificial (RNA) para predecir el rendimiento académico estudiantil. Las RNAs emulan el funcionamiento fisiológico del cerebro humano, tienen la capacidad de procesar y abstraer información y son empleadas en investigaciones relacionadas con modelado predictivo debido a su capacidad para identificar relaciones no lineales entre variables. Se emplea una base de datos con información académica, demográfica, social e institucional de 395 estudiantes colombianos de media vocacional de la Institución Educativa Villa del Socorro, Medellín (Colombia). La base de datos es construida mediante la aplicación de encuestas e informes institucionales antes del inicio de la pandemia COVID-19. Los resultados muestran que la RNA desarrollada aquí clasifica adecuadamente el 73% de la muestra y que tiene un mejor desempeño en métricas (accuracy, recall, precision y F1-Score) que otras técnicas de aprendizaje supervisado. Se concluye que la predicción temprana del rendimiento académico permite formular estrategias didácticas y pedagógicas que hacen más eficiente el proceso de enseñanza y aprendizaje.

Palabras clave: inteligencia artificial; RNA; educación; rendimiento; estudiantes

Development of an artificial neural network for predicting student academic success

Abstract

This study presents an artificial neural network (ANN) that is designed for predicting student academic performance. ANNs emulate human brain physiological function, have the ability to process information, and are used in research for making prediction models because of their ability to identify nonlinear relationships between variables. The database used contains academic, demographic, social, and institutional information of 395 Colombian vocational high school students at the Villa del Socorro Educational Institution, Medellín (Colombia). This information is retrieved from application surveys and institutional reports completed before the onset of the COVID-19 pandemic. The results show that the ANN developed here adequately classifies 73% of the sample and performs better in metrics (accuracy, recall, precision, and F1-Score) than other supervised learning techniques. It is concluded that early prediction of academic performance allows formulating didactic and pedagogical strategies that improve the efficiency of teaching and learning processes.

Keywords: artificial intelligence; ANN; education; performance; students

INTRODUCCIÓN

El capital humano y la educación son dos conceptos que están directamente relacionados. El rendimiento académico de un estudiante refleja el rendimiento que tendrá en la vida laboral (Deng et al. 2020), diversas investigaciones han estudiado los factores que influyen en el éxito o el fracaso escolar (Lindqvist y Vestman, 2011). Factores personales, familiares, económicos e institucionales pueden influir en el rendimiento académico estudiantil (Britton y Propper, 2016; Khanam y Nghiem, 2016). Estas investigaciones muestran que la autoeficacia, la automotivación y la autonomía están directamente relacionadas con el rendimiento académico estudiantil (Byura, 1977; Maher et al., 2018). Por lo general, los estudiantes que tienen dificultades de aprendizaje independientemente del nivel de motivación y conocimientos previos que posean antes del inicio de un curso, tienen dificultades para retener información en su memoria de trabajo, responden de manera impulsiva y desinhibida (Cortez y Silva 2008), tienen dificultades para representar conceptos mentalmente y su sentido está poco desarrollado (deRegnier, 2018). En Colombia, la predicción del rendimiento académico estudiantil en colegios tiene un bajo nivel de implementación, lo que obedece principalmente al desconocimiento y a la falta de habilidades por parte de los docentes y directivos (Barrera-Osorio, 2020). Por su parte, en este país una baja cantidad de universidades e Instituciones de Educación Superior (IES) emplean modelos predictivos enfocados a detectar la deserción estudiantil (Guarín, 2015).

En este trabajo se entrena una RNA, específicamente un perceptrón multicapa (MLP) para predecir el éxito o el fracaso estudiantil, este permite clasificar a los estudiantes que ingresan a un curso en un estado (Aprobado=1 / Reprobado=0). Esta clasificación es útil en la personalización de estrategias didácticas y pedagógicas de los docentes al momento de iniciar los cursos, lo que contribuye a una mayor eficiencia del proceso de enseñanza y a un mejor rendimientoacadémico de los estudiantes. Los docentes al conocer las principales características de los estudiantes, hábitos de estudio, aspectos familiares y personales, entre otros, pueden implementar estrategias con anticipación, personalizar los métodos de enseñanza, identificar los estilos de aprendizaje de los estudiantes y remitir a unidades de apoyo en caso de ser necesario. Para el entrenamiento de la RNA (MLP) se empleó Google Colaboratory, específicamente la función MLPRegressor, los hiperparámetros que optimizaban los resultados del entrenamiento fueron identificados haciendo uso de GridSearchCV, clase disponible en Scikit-Learn que permite evaluar y seleccionar de forma sistemática los parámetros de un modelo de clasificación o regresión. Los datos fueron normalizados antes del entrenamiento con la función MinMaxScaler. En resumen, se empleó una sola capa oculta con 9 neuronas, la función de activación logística, una tasa de aprendizaje adaptativa de 0.01, un momentum 0.9 y 1000 iteraciones. El 75% de los datos (n = 296) de la muestra fueron empleados para el entrenamiento y el 25% restante (n = 99) para la validación. Se empleó una RNA para la predicción porque permite incluir variables difíciles de cuantificar a través de ecuaciones simples, permitiendo considerar información personal de los estudiantes. Finalmente, se muestra que la RNA (MLP) desarrollada es adecuada para identificar a los estudiantes que aprobarán o reprobarán y se hace un comparativo con otros modelos de aprendizaje supervisado.

OTROS ANTECEDENTES

Las RNA son redes informáticas que tienen como objetivo emular el proceso de toma de decisiones de las neuronas en el sistema nervioso central biológico de los humanos o animales. Estas consisten en una simulación deneurona a neurona, que toma como base la neurofisiología, por lo tanto, difiere de las máquinas informáticas digitales y analógicas que se utilizan para mejorar o acelerar el cómputo del cerebro humano sin tener en cuenta la organización de los elementos informáticos y su conexión en red (Agudelo et al., 2014). Las RNA son herramientas poderosas para resolver problemas complejos relacionados con diversos usos, lo que reduce significativamente el tiempo y los gastos. Estas han sido utilizadas por diferentes investigadores para hacer predicciones (Ghazvini et al., 2020, Villada et al., 2016). La Figura 1 presenta la estructura de una neurona biológica.

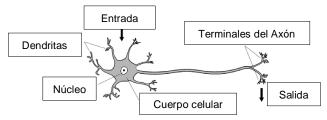
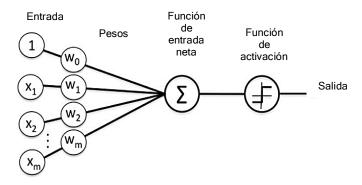


Fig. 1: Neurona biológica. Adaptado de (Neves et al. 2017)

El MLP tiene un mayor uso que los otros tipos de RNA, este requiere una estructura predefinida para su entrenamiento dependiendo del objetivo de predicción (Bhattacharjee y Pant,2019, Ramos et al., 2016). El MLP contiene mínimo tres capas (entrada, oculta y salida), estos tienen la capacidad de realizar un ajuste no lineal con alta precisión, su objetivo es encontrar valores adecuados para los pesos de los enlaces entre las neuronas, la técnica de entrenamiento más utilizada en este tipo de RNA es la retropropagación (Zarei y Behyad 2019). En Figura 2 se presenta la estructura general de un MLP.



Fig, 2: Perceptrón Multicapa. Fuente: Adaptado de (Nicholson 2019)

Generalmente, un MLP de una capa oculta es una función $f: R^E \to R^M$, donde E es el tamaño del vector de entrada x y M es el tamaño del vector de salida f(x) como se presenta en la ecuación 1:

$$f(x) = G(b^2 + W^2 + (S(b^1 + W^1 x)))$$
(1)

Donde b^1 y b^2 son los vectores de sesgo, W^1 y W^2 las matrices de peso, G y S las funciones de activación (Bhattacharjee y Pant 2019). El vector presentado en la ecuación 2 constituye la capa oculta:

$$h(x) = \emptyset(x) = S(b^1 + W^1 x)))$$
(2)

Donde $W^1 \in \mathbb{R}^{D \times Dh}$ es la matriz de peso que conecta el vector de entrada con la capa oculta. Cada columna W^1i representa los pesos de las unidades de entrada a *la i-ésima* unidad oculta. Las opciones típicas para S puede ser la función tangente hiperbólica (ecuación 3) o la función logística (ecuación 4).

$$tanh(x) = \frac{(e^x - e^{-x})}{(e^x + e^{-x})}$$
(3)

$$S(x) = \frac{1}{1 + e^{-x}} \tag{4}$$

En este trabajo se utilizó la función logística porque permite un entrenamiento más rápido y mejores mínimos locales. Tanto la función tangencial como la función logística son escalares, pero su extensión natural a los vectores y tensores se aplican en función de los elementos (Chai et al. 2017). El vector de salida se obtiene como se presenta en la ecuación 5:

$$f(x) = G(b^2 + W^2 + (S(b^1 + W^1 x)))$$
(5)

La probabilidad de pertenecer a una clase puede obtenerse eligiendo G como función de máxima seguridad. Para entrenar una RNA-MLP se entrenan todos los parámetros delmodelo Y, y se utiliza el Descenso de Gradiente Estocástico con *minibatches*. El conjunto de parámetros a aprender es el conjunto $\theta = \{W^2, b^2, W^2, b^3\}$ obteniendo los gradientes $\delta I / \delta \theta$ (Dai, Ma, y Xie 2014). La gráfica de la función logística construida sobre el plano cartesiano (X vs Y), tiene forma de S como se presenta en la Figura S. Esta función es útil cuando la salida que se va a predecir es binaria (S). Por lo general, se emplea en modelos que tienen como objetivo predecir probabilidades. Además, esta función es derivable, lo que permite encontrar la pendiente de la curva en dos puntos (Sharma 2017).

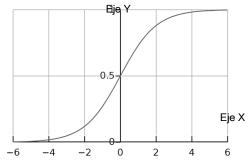


Fig. 3 Función Logística. Adaptado de Sharma (2017)

Diferentes autores han empleado RNA en el campo de la educación. Gorr et al., (1994) desarrollaron un estudio comparativo de RNA y modelos estadísticos para predecir promedios de calificaciones. Moucary et al., (2011) desarrollaron un procedimiento híbrido basado en RNA y en la puesta en común de datos que permite predecir el promedio académico (GPA) estudiantil según su rendimiento. Naik y Ragothaman (2004) emplearon RNA para predecir el éxito de los estudiantes de MBA y demostraron su buen funcionamiento. Herzog (2006) estimó el tiempo de retención y el tiempo de finalización de los estudiantes universitarios. Stevens et al., (1996), docentes del área de la medicina, identificaron estrategias expertas para la resolución de problemas de salud a partir de un banco de información de diagnósticos de pacientes haciendo uso de RNA.

METODOLOGÍA

Para la construcción de la RNA-MLP se empleó Google Colab, plataforma web de acceso gratuito que permite emplear cuadernos de Jupyter, combinar código ejecutable y texto, imágenes, HTML y LaTeX. Es compatible con librerías de Python para el análisis y visualización de datos como Pandas, Numpy, Matplotlib, Keras, Tensorflow, entre otras, los cuales se almacenan en una cuenta de Google Drive. Como se mencionó anteriormente, se empleó un conjunto de datos de 395 estudiantes colombianos (243 hombres, Edad Media = 17.39 años, Desviación Estándar (SD) = 1.24 años / 133 mujeres, Edad Media = 16.79 años, Desviación Estándar (SD) = 1.23 años) de décimo y undécimo grado de formación con información académica, demográfica, social e institucional de la Institución Educativa Villa del Socorro, Medellín, Colombia, la cual fue recogida mediante la aplicación de encuestas e informes institucionales durante el primer trimestre del año 2020, antes del inicio de la pandemia generada por el COVID-19 en Colombia. Toda la población estudiantil matriculada en esos dos grados fue considerada en el estudio (Censo), por lo cual no se realizó una selección del tamaño de la muestra. En total se analizaron 29 variables. Se empleó la función MLPRegressor, se identificaron los parámetros óptimos con GridSearchCV, se normalizaron los datos con la función MinMaxScaler, se utilizó una sola capa oculta con 9 neuronas, la función de activación logística, una tasa de aprendizaje adaptativa inicial de 0.01, un momentum 0.9 y 1000 iteraciones (Ver Figura 4). El 75% de los datos (n = 296) de la muestra fueron empleados para el entrenamiento y el 25% restante (n = 99) para la validación. La Tabla 1 presenta la descripción de las variables de la base de datos.

```
from sklearn.neural_network import MLPRegressor

m = MLPRegressor(
   hidden_layer_sizes=(9,),
   activation="logistic",
   learning_rate="adaptive",
   momentum=0.9,
   learning_rate_init=0.01,
   max_iter=1000,
)
```

Fig. 4 Código-MLP. Fuente: elaboración de los autores

Tabla 1: Descripción	de las variables
----------------------	------------------

Nombre	Descripción			
Var1-Sexo	Masculino / Femenino			
Var2-Edad	Numero entero entre 14 y 21 años			
Var3-Procedencia	Rural / Urbana			
Var4-Tamaño Familiar	≤ 3 / > 3			
Var5-Habita con los padres	Si / No			
Var6-Nivel educativo de la madre				
Var7- Nivel educativo del padre	Ninguno / primaria / secundaria / técnico / universitario			
Var8-Trabajo de la madre	Educación / Salud / Civil / Hagar / Otro			
Var9-Trabajo del padre	Educación / Salud / Civil / Hogar / Otro			
Var10-Razón de elección el colegio	Cercanía al hogar / reputación del colegio / preferencia / otro			
Var11-Acudiente del estudiante	Madre / Padre / Otro			
Var12-Tiempo de traslado de la casa al colegio	1 - <15 min / 15 a 30 min / 30 min a 1 hora / 4 - >1 hora			
Var13-Tiempo de estudio semanal	<2 horas / 2 a 5 horas/ 5 a 10 horas / >10 horas			
Var14-Número de fallas	Número entero			

Tabla 1: continuación

Nombre	Descripción			
Var15-Soporte educativo extra				
Var16-Soporte educativo familiar				
Var17-Paga clases extra				
Var18-Actividades extra-curriculares				
Var19-Acude a apoyo en el colegio	Si = 0 / No =1			
Var20-Desea entrar a la Universidad				
Var21-Acceso a internet en casa				
Var22-Relación sentimental				
Var23-Calidad de las relaciones familiares	Muy mala / Mala / Media / Buena / Excelente			
Var24-Tiempo uso redes sociales/día	1 - <15 min / 15 a 30 min / 30 min a 1 hora / 4 - >1 hora			
Var25-Salidas con amigos	Si = 0 / No =1			
Var26-Consumo de alcohol diario				
Var27-Consumo de alcohol los fines de semana				
Var28-Estado actual de salud	Muy mala / Mala / Media / Buena / Excelente			
Var29-Número de ausencias al colegio	Número entero			
Var30-Aprobado	Si = 0 / No =1			

RESULTADOS Y DISCUSIÓN

En esta sección se presenta la información de los datos empleados y el análisis realizado sobre los mismos. La Tabla 2 presenta el resumen de los datos de la RNA, 296 estudiantes fueron asignados a la muestra de entrenamiento y 99 a la muestra de prueba. Todos los estudiantes (395) fueron incluidos en el análisis. La Tabla 3 presenta la información de cada una de las capas de la RNA.

Tabla 2: Resumen de las muestras

		Ν	%
Muestra	Entrenamiento	296	75
Muestra	Prueba	99	2
	Muestra	395	100
Total	Excluido	0	0
	Total	395	100

Tabla 3: Resumen de las muestras

	Factores	V1, V2				
Capa de entrada	Co-variables	V3, V4, V5, V6, V7, V8, V9, V10, V11, V12, V13, V14, V15, V16, V17, V18, V19, V20, V21, V22, V23, V24, V25, V26, V27, V28, V29				
Capa oculta	Número de unidades	Corrección estandarizada – Capa oculta (9).				
	Función de activación	Logística				
Capa de salida	Variable dependiente	Aprobado/Reprobado				
	Número de unidades	1				
	Función de error	Suma de cuadrados				

La Tabla 4 presenta la información resumida sobre los resultados del entrenamiento y prueba, se presenta la suma del error cuadrático medio (ECM), este es el error que la RNA trata de minimizar durante el entrenamiento y la prueba. Sea \hat{Y} un vector de n predicciones y un vector de valores verdaderos, el ECM es definido como:

$$ECM = \frac{1}{n} \sum_{i=1}^{n} (\hat{Y}_i - Y_i)^2$$
 (6)

Mientras más se acerque a cero el valor del *ECM* menor es la cantidad de error entre los dos conjuntos de datos. El ECM de la muestra de prueba (51,678) es inferior al ECM de la muestra de entrenamiento (94,362). La Tabla 4 presenta la matriz de confusión de la muestra de entrenamiento (*n*=99). La respuesta esperada en cada casoes la categoría (Aprobado/Reprobado) con mayor probabilidad asignada por el modelo, el modelo clasificó correctamente el 73% de los casos. Adicionalmente, se entrenaron los modelos de clasificación; Naive Bayes, Regresión Logística, Support Vector Machine (SVM), K-Nearest Neigbors (KNN), Random Forests y Árboles de Decisión en Google Colab con la misma base de datos normalizada de los estudiantes. La Fig.5 presenta las librerías de Python empleadas para entrenar los modelos.

Tabla 4: Matriz de confusión

	Predicción				
ual		Aprueba	Reprueba		
Act	Aprueba	53	9		
	Reprueba	18	19		

from sklearn.naive_bayes import GaussianNB
from sklearn.linear_model import LogisticRegression
from sklearn.svm import SVC
from sklearn.neighbors import KNeighborsClassifier
from sklearn.ensemble import RandomForestClassifier
from sklearn.tree import DecisionTreeClassifier

Fig. 5 Librerías - Modelos de clasificación. Fuente: elaboración de los autores

El objetivo de entrenar estos modelos fue realizar un comparativo de las métricas; Accuracy, Recall, Precision y F1-Score, con las obtenidas por la RNA. Es necesario aclarar que, el Accuracy mide cuántas observaciones positivas y negativas se clasificaron correctamente y es una métrica adecuada para comparar y seleccionar diferentes modelos de clasificación o regresión, siempre y cuando se tengan conjuntos de datos simétricos en los cuales los valores de falsos positivos (FP) y falsos negativos (FN) sean equivalentes. El Accuracy se calcula como la suma del número total de verdaderos positivos (TP) y el total de verdaderos negativos (TN) entre la suma del número total de verdaderos positivos (TP), falsos negativos (FN), verdaderos negativos (TN) y falsos positivos (FP). Recall es el número de verdaderos positivos (TP) dividido entre el número de verdaderos positivos (TP) y falsos negativos (FN) en los datos de prueba, la precisión es el número de verdaderos positivos (TP) dividido entre todas las predicciones positivas (TP y FP), una precisión alta se relaciona con una tasa baja de falsos positivos (FP). Por último, el F1-Score es la media armónica entre Precisión y Recall, esta métrica es especialmente útil si se tiene una distribución de clases desigual. A continuación, se presentan las ecuaciones de cada una de las métricas mencionadas anteriormente:

$$Accuracy = \frac{TP + TN}{TP + FN + TN + FP} \tag{7}$$

$$Recall = \frac{TP}{TP + FN} \tag{8}$$

$$Precision = \frac{TP}{TP + FP} \tag{9}$$

$$F1 - Score = 2 \frac{Precision.Recall}{Precision + Recall}$$
(10)

La Tabla 5 presenta un comparativo general de los resultados de las métricas analizadas, se puede observar que la RNA-MLP es el modelo que registra los mejores resultados. Considerando que la desviación estándar (σ^2) de las métricas analizadas en cada uno de los modelos es cercana a cero, es decir, no se alejan entre sí, la Media (\bar{x}) es la Mediad de Tendencia Central (MTC) más adecuada para analizar hacían donde se tienden a concentrar los datos. Se puede apreciar que todas las métricas obtenidas por la RNA son superiores a la Media (\bar{x}), lo que indica una vez más que es el modelo más adecuado para predecir el rendimiento académico de los estudiantes. La Tabla 6 presenta las tres primeras variables en orden de importancia según la clasificación de la RNA-MLP. El tiempo de estudio y las ausencias son los factores más significativos para predecir si un estudiante aprobará o reprobará, independientemente de su sexo, edad y demás variables consideradas en la base de datos. Se debe aclarar que, los resultados no muestran otros factores como significativos porque la población era muy homogénea en cuanto a otras variables que no resultaron significativas.

Tabla 5: Comparativo general

	Naive Bayes	Regresión Logística	SVM	KNN	Random Forests	Árboles de Decisión	RNA	Media (X̄)	Desviación Estándar (σ²)
Accuracy	0,68	0,69	0,71	0,67	0,71	0,69	0,73	0,70	0,02
F1-Score	0,69	0,70	0,72	0,68	0,72	0,68	0,74	0,70	0,02
Recall	0,68	0,71	0,73	0,67	0,72	0,69	0,74	0,71	0,03
Precision	0,68	0,71	0,73	0,67	0,72	0,69	0,74	0,71	0,03

Tabla 6: Importancia de las variables

Variable	Importancia relativa %	Orden
Tiempo de estudio	18.7%	1
Ausencias	17.6%	2
Tiempo de uso redes sociales / día	16.2 %	3

Diferentes investigaciones coinciden con los resultados de la RNA-MLP desarrollada. La asociación entre la calidad del sueño y el rendimiento académico en adolescentes está mediada por el tiempo de uso de redes sociales, reducir su uso puede ser una intervención factible para mejorar la calidad del sueño, lo que genera efectos potencialmente positivos sobre el rendimiento académico (Adelantado-Renau et al., 2019). Por otra parte, Plant et al., (2005), afirman que la cantidad de tiempo individual de estudio es un predictor potencial del éxito académico, a mayor cantidad de tiempo mejores calificaciones, sin dejar de lado la importancia de fortalecer las habilidades personales. Por su parte, Marta et al., (2020), identificaron factores biológicos y psicosociales relacionados con el rendimiento académico, incluidos los cambios hormonales en las mujeres, el síndrome de la fase de sueño tardía, la interacción con los compañeros en la vida académica y social, trabajos y actividades extracurriculares, emociones negativas, consumo de tabaco, alcohol y café, problemas familiares y uso desmedido de internet y dispositivos electrónicos. Xie et al., (2020), señalan que la motivación académica es un factor importante en el rendimiento académico de los estudiantes, identificaron una relación entre la motivación y el compromiso de los estudiantes en conductas de orientación académica, como estudiar, participar en discusiones en clase, trabajar con compañeros y realizar cursos más exigentes, estas actividades conducen directamente a un mayor aprendizaje, a un mayor rendimiento académico y a una mejor preparación para el éxito universitario. Finalmente, Begdache et al., (2020) señalan la importancia que tiene la salud mental sobre el rendimiento académico de los estudiantes; la angustia mental baja está asociada positivamente con buenos logros académicos y somnolencia diurna limitada, la angustia mental leve está relacionada con la negligencia académica y con un promedio de calificaciones bajo y la angustia mental severa está relacionada con somnolencia diurna excesiva y bajo rendimiento académico. Los estilos de vida de los estudiantes contribuyen a la salud mental de los mismos, los cuales se convierten en eventos potencialmente cíclicos que pueden afectar su rendimiento académico.

CONCLUSIONES

En esta investigación se entrenó una RNA para predecir el rendimiento académico estudiantil. La RNA clasificó adecuadamente el 73% de la muestra de prueba, lo que nos permite decir que la RNA es correcta para identificar qué estudiantes aprobarían o no. Esta información puede ayudar a los profesores a tomar medidas preventivas para evitar el fracaso escolar de los estudiantes, también puede ayudarles a emplear estrategias personalizadas a los estudiantes que tienen problemas económicos, familiares o de otro tipo y remitirlos a las unidades de permanencia académica o afines. El tiempo de estudio, las ausencias y el tiempo de uso de redes sociales son los factores más importantes para determinar la probabilidad de que un estudiante apruebe o no un curso. La clasificación de los estudiantes según sus condiciones familiares, sociales y económicas puede ayudar a identificar a los posibles estudiantes que tendrán un rendimiento académico inferior al esperado, ya sea por malos hábitos de estudio, problemas económicos en el hogar, malas relaciones con los miembros de la familia o con personas cercanas a ellos.

REFERENCIAS

Adelantado-Renau, M., Diez-Fernandez, A., y tres cuatro autores, The effect of sleep quality on academic performance is mediated by Internet use time: DADOS study, doi.org/10.1016/j.jped.2018.03.006, J. de Pediatría, 95(4), 410-418 (2019)

Agudelo, L., Velilla, E., y López, J. M., Estimación de la carga de transformadores de potencia utilizando una red neuronal artificial, doi.org/ 10.4067/S0718-07642014000200003, Información Tecnológica, 25(2), 15-23 (2014)

Barrera-Osorio, F., Gonzalez, K., Lagos, F., y Deming, D. J., Providing performance information in education: an experimental evaluation in Colombia, doi.org/10.1016/j.jpubeco.2020.104185, Journal of Public Economics, 186, 104185 (2020)

Begdache, L., Kianmehr, H., y otros tres autores, Principal component regression of academic performance, substance use and sleep quality in relation to risk of anxiety and depression in young adults, doi.org/10.1016/j.tine.2019.03.002, Trends in Neuroscience and Education, 15, 29-37 (2019)

Bhattacharjee, K., y Pant, M., Hybrid particle swarm optimization-genetic algorithm trained multi- layer perceptron for classification of human glioma from molecular brain neoplasia data, doi.org/10.1016/j.cogsys.2019.06.003, Cognitive Systems Research, 58, 173–94 (2019)

Britton, J., y Propper, C., Teacher pay and school productivity: exploiting wage regulation, doi.org/10.1016/j.jpubeco.2015.12.004, J. of Public Economics, 133, 75–89 (2016)

- Chai, M., Luo, K., Shao, C., y Fan, J., An efficient level set remedy approach for simulations of two-phase flow based on sigmoid function, doi.org/10.1016/j.ces.2017.06.009, Chemical Engineering Science, 172, 335–52 (2017)
- Cortez, P., y Silva, A. M. G., Using data mining to predict secondary school student performance. In 15th European Concurrent Engineering Conference 2008, ECEC 2008 5th Future Business Technology Conference, http://www3.dsi.uminho.pt/pcortez/, FUBUTEC 2008, 5–12 (2008)
- Dai, Q., Ma, Z., y Xie, Q., A two-phased and ensemble scheme integrated backpropagation algorithm, doi.org/10.1016/j.asoc.2014.08.012, Applied Soft Computing Journal, 24, 1124–35 (2014)
- Deng, W., Li, X., Wu, H., y Xu, G., Student leadership and academic performance, doi.org/10.1016/j.chieco.2019.101389, China Economic Review, 60, 101389 (2020)
- deRegnier, R. A., One step forward for evidence-based therapy for learning problems in preterm children, doi.org/10.1016/j.jpeds.2018.09.014, The Journal of Pediatrics, 202,1 (2018)
- Ghazvini, M., Maddah, H., y otros tres autores, Experimental evaluation and artificial neural network modeling of thermal conductivity of water based nanofluid containing magnetic copper nanoparticles, doi.org/10.1016/j.physa.2019.124127, Physica A: Statistical Mechanics and Its Applications, 24, 124127 (2020)
- Gorr, W. L., Nagin, D., y Szczypula, J., Comparative study of artificial neural network and statistical models for predicting student grade point averages, doi.org/10.1016/0169-2070(94)90046-9, International Journal of Forecasting 10 (1), 17–34 (1994)
- Guarín, C. E. L., Guzmán, E. L., y González, F. A., A model to predict low academic performance at a specific enrollment using data mining, doi.org/10.1109/RITA.2015.2452632, IEEE Revista Iberoamericana de Tecnologías del Aprendizaje, 10(3), 119-125 (2015)
- Herzog, S., Estimating student retention and degree-completion time: decision trees and neural networks vis-à-vis regression, doi.org/10.1002/ir.185, New Directions for Institutional Research, 131, 17–33 (2006)
- Khanam, R., y Nghiem, S., Family income and child cognitive and noncognitive development in Australia: does money matter?, doi.org/10.1007/s13524-016-0466-x, Demography, 53 (3), 597–621 (2016)
- Lindqvist, E., y Vestman, R., The labor market returns to cognitive and noncognitive ability: evidence from the Swedish Enlistment, American Economic Journal: Applied Economics, 3(1), 101–28 (2011)
- Maher, C. A., Sigley, R., Sullivan, P., y Wilkinson, L. C., An international perspective on knowledge in teaching mathematics, doi.org/10.1016/j.jmathb.2018.05.002, J. of Mathematical Behavior, 51, 71–79 (2018)
- Marta, O. F. D., Kuo S.Y., y otros cinco autores, Gender differences in the relationships between sleep disturbances and academic performance among nursing students: a cross-sectional study, doi.org/10.1016/j.nedt.2019.104270, Nurse Education Today, 85, 104270 (2020)
- Moucary, C. E., Khair, M., y Zakhem, W., Improving student's performance using data clustering and neural networks in foreign-language based higher education, The Research Bulletin of Jordan ACM, 2(3), 27–34 (2011)
- Bijayanya, N., y Ragothaman, S., Using neural networks to predict MBA student success, College Student Journal, 38(1), 143–50 (2004)
- Neves, A. C., González, I., Leander, J., y Karoumi, R., A new approach to damage detection in bridges using machine learning, doi.org/10.1007/978-3-319-67443-8_5, In International Conference on Experimental Vibration Analysis for Civil Engineering Structures, 73–84, Springer (2017)
- Nicholson, Ch., A beginner's guide to neural networks and deep learning, A.I WiKi, https://wiki.pathmind.com/neural-network/ (2019)
- Plant, E. A., Ericsson, K. A., Hill, L., y Asberg, K., Why study time does not predict grade point average across college students: implications of deliberate practice for academic performance, doi.org/10.1016/j.cedpsych.2004.06.001, Contemporary Educational Psychology, 30(1), 96-116 (2005)
- Ramos, O. L., Rojas, D. A., y Saby, J. E., Reconocimiento de patrones vocálicos mediante la implementación de una red neuronal artificial utilizando sistemas embebidos, doi.org/10.4067/S0718-07642016000500015, Información Tecnológica, 27(5), 133-142 (2016)
- Sharma, S., What is activation function? Towards data science, https://towardsdatascience.com/ (2017)
- Stevens, R. H., Lopo, A. C., y Wang, P., Artificial neural networks can distinguish novice and expert strategies during complex problem solving, doi.org/10.1136/jamia.1996.96236281, J. of the American Medical Informatics Association, 3(2), 131–38 (1996)
- Villada, F., Muñoz, N., y García-Quintero, E., Redes neuronales artificiales aplicadas a la predicción del precio del oro, doi.org/10.4067/S0718-07642016000500016, Información Tecnológica, 27(5), 143-150 (2016)
- Xie, K., Vongkulluksn, V. W., Lu, L., y Cheng, S. L., A person-centered approach to examining high-school students' motivation, engagement and academic performance, doi.org/10.1016/j.cedpsych.2020.101877, Contemporary Educational Psychology, 62,101877-101890 (2020)
- Zarei, T., y Behyad, R., Predicting the water production of a solar seawater greenhouse desalination unit using multi-layer perceptron model, doi.org/10.1016/j.solener.2018.11.059, Solar Energy, 177, 595–603 (2019)