

<https://doi.org/10.23913/ride.v12i24.1180>

Artículos científicos

Algoritmos de aprendizaje automático para la predicción del logro académico

Machine Learning Algorithms for Predicting of Academic Achievement

*Algoritmos de aprendizaje de máquina para previsão de desempenho
acadêmico*

Miguel Ángel Morales Hernández

Colegio de Postgraduados, Campus Montecillo, México

morales.miguel@colpos.mx

<https://orcid.org/0000-0002-0351-9356>

Juan Manuel González Camacho

Colegio de Postgraduados, Campus Montecillo, México

jmgc@colpos.mx

<https://orcid.org/0000-0001-5479-7316>

Héctor Robles Vásquez

Planeación y Evaluación del Consejo Nacional de Fomento Educativo, México

hroblesvasquez@gmail.com

<https://orcid.org/0000-0001-9759-309X>

David H. del Valle Paniagua

Colegio de Postgraduados, Campus Montecillo, México

dhvallep@colpos.mx

<https://orcid.org/0000-0003-4383-4323>

José Rafael Durán Moreno

Consultor independiente, México

rduran1091@gmail.com

<https://orcid.org/0000-0002-7886-6408>



Resumen

En esta investigación se implementaron dos clasificadores de aprendizaje automático, una red neuronal multicapa (perceptrón multicapa [MLP]) y un modelo de potenciación del gradiente (GB), para predecir el grado de logro académico en las asignaturas de español y matemáticas de alumnos de sexto de primaria (2008) y tercero de secundaria (2011) con base en variables contextuales obtenidas de los Exámenes Nacionales del Logro Académico en Centros Escolares (Enlace) del estado de Tlaxcala, México. Se consideraron 13 variables de entrada y la importancia relativa de éstas, se determinó por medio del algoritmo bosque aleatorio (RF). Los clasificadores MLP y GB se entrenaron y probaron con un conjunto de datos de 11 036 registros de estudiantes que permanecieron en el sistema escolar de 2008 a 2011. Los modelos se entrenaron y probaron en predicción para 2008 y 2011. En español MLP fue superior a GB con una precisión global de clasificación (PG) de 70.1 % en 2008 y 61.1 % en 2011. GB obtuvo mejores resultados en matemáticas con una PG de 68.8 % en 2008 y 63.5 % en 2011. Se observó que el puntaje en español tiene una fuerte asociación con el grado de logro académico en matemáticas. Los puntajes en español y matemáticas tuvieron mayor importancia relativa con respecto a los factores contextuales analizados como: sexo, beca, turno de la escuela. En la población de alumnos analizada se observó que en español y matemáticas la proporción de mujeres es mayor a la proporción de hombres en los grados de logro académico elemental, y bueno o excelente; en contraste, en ambas asignaturas esta proporción se invierte con el grado de logro insuficiente.

Palabras clave: aprendizaje supervisado, árboles de decisión, contexto escolar, redes neuronales artificiales, validación cruzada.

Abstract

In this research, two machine learning classifiers were implemented, a multilayer perceptron (MLP) and a gradient boosting model (GB), to predict the degree of academic achievement in Spanish and mathematics of basic education students in two stages, sixth of primary (2008) and third of secondary (2011), based on contextual variables obtained from the Enlace test of the state of Tlaxcala, Mexico. Thirteen input variables were considered. The relative importance of these was determined by the random forest (RF) classifier. MLP and GB classifiers were trained and tested with a dataset of 11 036 records of students who remained

in the school system from 2008 to 2011. The models were trained and tested in prediction for 2008 and 2011. In Spanish MLP outperformed GB with a global classification accuracy (PG) of 70.1 % in 2008 and 61.1 % in 2011. GB obtained better performance in mathematics with a PG of 68.8 % in 2008 and 63.5 % in 2011. It was observed that the score in Spanish has a strong association with the degree of academic achievement in mathematics. Scores in Spanish and mathematics have greater relative importance with respect to contextual factors considered as sex, scholarship, school shift, and so on. In the population of students analyzed, it is observed that, in Spanish and mathematics, the proportion of women is higher than the proportion of men in achievement levels 1 (elementary) and 2 (good or excellent); in contrast, in both subjects this proportion is reversed at achievement level 0 (insufficient).

Keywords: supervised learning, decision trees, school context, artificial neural networks, cross validation.

Resumo

Nesta pesquisa, dois classificadores de aprendizado de máquina, uma rede neural multicamada (multilayer perceptron [MLP]) e um modelo de potenciação de gradiente (GB), foram implementados para prever o grau de desempenho acadêmico nas disciplinas de espanhol e matemática de alunos do ensino médio. sexta série (2008) e terceira série (2011) com base em variáveis contextuais obtidas nos Exames Nacionais de Desempenho Acadêmico nas Escolas (Enlace) do estado de Tlaxcala, México. 13 variáveis de entrada foram consideradas e sua importância relativa foi determinada usando o algoritmo Random Forest (RF). Os classificadores MLP e GB foram treinados e testados com um conjunto de dados de 11.036 prontuários de alunos que permaneceram na rede escolar de 2008 a 2011. Os modelos foram treinados e testados em previsão para 2008 e 2011. Em espanhol, o MLP foi superior ao GB com uma precisão geral de notas (GP) de 70,1% em 2008 e 61,1% em 2011. GB teve um desempenho melhor em matemática com um GP de 68,8% em 2008 e 63,5% em 2011. A pontuação em espanhol mostrou ter uma forte associação com o grau de desempenho acadêmico em matemática. Os escores em espanhol e matemática tiveram maior importância relativa em relação aos fatores contextuais analisados como: gênero, escolaridade, turno escolar. Na população de alunos analisada, observou-se que em espanhol e matemática a proporção de mulheres é maior do que a proporção de homens no ensino

fundamental e nas notas de desempenho acadêmico bom ou excelente; e essa proporção se inverte com o grau de realização insuficiente.

Palavras-chave: aprendizagem supervisionada, árvores de decisão, contexto escolar, redes neurais artificiais, validação cruzada.

Fecha Recepción: Octubre 2021

Fecha Aceptación: Abril 2022

Introducción

La evaluación del aprendizaje de los alumnos mediante pruebas a gran escala (estatal o nacional) permite obtener información sobre su grado de logro académico y las variables contextuales asociadas. La Organización para la Cooperación y el Desarrollo Económicos [OCDE] (2005) encontró evidencias de cómo se relacionan factores como el contexto escolar, insumos y procesos escolares con el proceso de aprendizaje de los alumnos. México empezó a usar las pruebas estandarizadas para medir el logro académico de los estudiantes en las dos últimas décadas. La Secretaría de Educación Pública (SEP) cuenta con bases de datos de alumnos que se matriculan anualmente en cada nivel educativo y resultados de las pruebas que se aplican a nivel nacional como los Exámenes de Calidad y Logro Educativo (Excale), Evaluación Nacional de Logro Académico en Centros Escolares (Enlace) o a nivel internacional como el Programa Internacional de Evaluación de los Alumnos (PISA) (Instituto Nacional para la Evaluación de la Educación [INEE], 2019).

La prueba Enlace se aplicó de manera censal a partir de 2006 a todos los alumnos de tercero a sexto de primaria y de los tres años de secundaria. En 2008 se aplicó a los tres años de bachillerato. El objetivo de esta prueba fue evaluar el logro académico de los estudiantes en las asignaturas de español y matemáticas; el referente de esta prueba es el currículo nacional (Martínez, 2015). Los resultados de Enlace se miden en una escala estandarizada que va de 200 a 800 puntos y cuenta con cuatro niveles de logro (0 = Insuficiente, 1 = Elemental, 2 = Bueno, 3 = Excelente). La información ha servido de soporte a los docentes para comparar los resultados de su escuela con otras de características similares e identificar contenidos curriculares que los estudiantes no adquirieron para tomar acciones pertinentes (SEP, 2008). A la par de la prueba Enlace, se aplicaron cuestionarios a una muestra de alumnos, padres de familia, docentes y directores de las escuelas que fueron incluidas en la prueba para conocer las características personales, entorno familiar, hábitos de lectura,

características de la vivienda, infraestructura escolar y métodos de enseñanza, e identificar factores limitantes asociados al aprendizaje.

En la literatura se reportan aplicaciones de los modelos de aprendizaje automático para evaluar el rendimiento escolar. Đambić, Krajcar y Bele (2016) aplicaron un modelo de regresión logística para la detección temprana de estudiantes con problemas de rendimiento en un curso de computación. El modelo obtuvo un error de clasificación de 19.0 %. Rai *et al.* (2020) utilizaron dos modelos de clasificación (bosque aleatorio y máquina de soporte vectorial) para predecir el rendimiento escolar de un grupo de estudiantes universitarios con base en variables de entrada como: sexo, horas de estudio, porcentaje de asistencia a clases, e ingreso mensual de la familia. El modelo bosque aleatorio obtuvo una precisión global de clasificación de 94.0 % y la máquina de soporte vectorial de 79.0 %. Por su parte, Altabrawee, Osama y Qaisir (2019) aplicaron cuatro algoritmos de aprendizaje automático (red neuronal, decisión Bayesiana, árboles de decisión y regresión logística) para predecir el rendimiento de los estudiantes en un curso de computación con base en el uso del Internet como medio de aprendizaje; algunas variables consideradas fueron el tiempo empleado en redes sociales, horas de estudio, sexo, educación de los miembros de la familia. La red neuronal alcanzó el mejor desempeño, con una precisión global de 77.0 % y área bajo la curva AUC_{ROC} de 0.807.

En México se han realizado trabajos relacionados para evaluar el efecto de factores externos e internos a la escuela en el logro académico de los estudiantes. Fernández (2003) utilizó índices de capital familiar global (nivel educativo de la madre, equipamiento de confort en la vivienda, disponibilidad de libros y de computadora), de contexto sociocultural de la escuela y de clima organizacional; el análisis jerárquico se aplicó para evaluar el aprendizaje en español y en matemáticas. Este autor reporta que un aumento en el índice de capital familiar global incide en el incremento del desempeño en español y matemáticas; sin embargo, cuando se generalizan las carencias de la vivienda (analfabetismo, bajos ingresos) los resultados de los alumnos son bajos.

En esta investigación se plantearon los siguientes objetivos, en primer lugar, implementar dos clasificadores de aprendizaje automático supervisado, a saber, una red neuronal multicapa y un algoritmo de potenciación del gradiente, para predecir el grado de logro académico (0: insuficiente, 1: elemental, 2: bueno o excelente) en las asignaturas de español y matemáticas de alumnos de sexto de primaria (2008) y tercero de secundaria (2011)

del estado de Tlaxcala a partir de datos de la prueba Enlace; en segundo lugar, comparar el grado de logro académico en español y matemáticas en 2008 y 2011; y en tercero, determinar la importancia relativa de 13 variables predictoras en la clasificación del logro académico. Las variables predictoras fueron puntaje en matemáticas, puntaje en español, beca, turno de la escuela, sostenimiento, tipo de localidad, sexo, tipo de escuela, tamaño de localidad, marginación, ubicación geográfica (altitud, latitud y longitud).

Materiales y métodos

Colecta y preparación de datos

La base de datos de registros académicos utilizados en este estudio (2008-2011) para el estado de Tlaxcala corresponde a un subconjunto de la base datos nacional de la prueba Enlace que se aplicó de 2006 a 2014 a todos los alumnos de tercero de primaria a tercero de bachillerato, cuyo propósito fue generar información para los padres de familia, docentes, directivos y a la sociedad en general sobre el logro académico de los estudiantes del sistema educativo en español y matemáticas (SEP, 2008). Los registros de datos se almacenaron por alumno, año de aplicación, nivel educativo y grado; con datos sobre logro académico de los alumnos (tabla 1), puntaje obtenido en la prueba (escala de 200 a 800), becario, turno, tipo de sostenimiento y localización geográfica de la escuela.

Tabla 1. Intervalos de puntajes para determinar la clase o nivel de logro académico en español y matemáticas

Grado y año	Español		
	0 [†]	1 [¶]	2 [§]
6.º 2008	≤413.85	(413.85, 581.62)	(581.62, 714.01)
7. ^{op} 2009	≤446.31	(446.31, 593.18)	(593.18, 735.69)
8. ^{oa} 2010	≤445.08	(445.08, 592.41)	(592.41, 735.35)
9. ^{ot} 2011	≤462.94	(462.94, 608.22)	(608.22, 749.18)
	Matemáticas		
6.º 2008	≤412.62	(412.62, 608.13)	(608.14, 735.70)
7. ^{op} 2009	≤507.27	(507.27, 634.85)	(634.85, 737.26)
8. ^{oa} 2010	≤505.39	(505.39, 634.05)	(364.05, 737.32)
9. ^{ot} 2011	≤525.99	(525.99, 657.03)	(657.03, 762.21)

[†]0 = Insuficiente; [¶]1 = Elemental; [§]2 = Bueno o excelente. ^p7.º = 1.º de secundaria, ^o8.º = 2.º de secundaria, ^{††}9.º = 3.º de secundaria.

Fuente: Prueba Enlace 2008-2013 (SEP, 2008)

De la información disponible para el estado de Tlaxcala se seleccionó el subconjunto de alumnos que realizaron la prueba Enlace durante cuatro años consecutivos (2008 a 2011) de sexto de primaria hasta tercero de secundaria; este periodo marca el inicio de una trayectoria educativa (el paso de primaria hasta llegar a tercero de secundaria). Las variables de contexto seleccionadas son: puntaje en español, puntaje en matemáticas, becario, turno de la escuela, tipo de sostenimiento, tipo de localidad, sexo, tipo de escuela, tamaño de localidad medida por sus habitantes, nivel de marginación, la localización geográfica (latitud, longitud y altitud) del municipio en donde se encuentra la escuela; así como los puntajes en español y matemáticas (Tabla 2).

Tabla 2. Variables contextuales y localización geográfica seleccionadas del Estado de Tlaxcala, México.

Variable	Descripción	Valores
<i>n_esp</i>	Logro académico en español	0 = Insuficiente; 1 = Elemental; 2 = Bueno o excelente
<i>n_mat</i>	Logro académico en matemáticas	0 = Insuficiente; 1 = Elemental; 2 = Bueno o excelente
<i>p_esp</i>	Puntaje en español	De 200 a 800 puntos
<i>p_mat</i>	Puntaje en matemáticas	De 200 a 800 puntos
<i>becario</i>	Condición de becario	0 = No becario, 1 = Becario
<i>turno</i>	Turno de la escuela	0 = Matutino, 1 = Vespertino
<i>t_sost</i>	Tipo de sostenimiento	0 = Público, 1 = Privado
<i>t_loc</i>	Tipo de la localidad	0 = Urbano, 1 = Rural
<i>sexo</i>	Sexo	0 = Hombre, 1 = Mujer
<i>t_esc</i>	Tipo de escuela	1 = General, 2 = Indígena, 3 = Conafe, 4 = Particular, 5 = Telesecundaria, 6 = Técnica
<i>t_loc</i>	Tamaño de la localidad	1 = menos de 100 habs., 2 = 100 a 249 habs., 3 = 250 a 499 habs., 4 = 500 a 2499 habs., 5 = 2500 a 14 999 habs., 6 = 15 000 ó más habs.
<i>n_mar</i>	Nivel de marginación	1 = Muy alto, 2 = Alto, 3 = Medio, 4 = Bajo, 5 = Muy bajo

<i>lat</i>	Latitud	Decimal
<i>lon</i>	Longitud	Decimal
<i>alt</i>	Altitud	Metros sobre el nivel del mar
<i>pob</i>	Población	En miles de habitantes

Fuente: Elaboración propia con base en los formatos F911 de la prueba Enlace (SEP, 2008) y con datos del Catálogo Único de Claves de Áreas Geoestadísticas Estatales, Municipales y Localidades (Inegi, 2020)

En la tabla 3 se presenta la descripción de los cuatro conjuntos de datos analizados por asignatura y año para la misma población escolar. Con el propósito de reducir el desbalance entre clases o niveles de logro académico los niveles bueno y excelente, se agruparon en el nivel dos de logro académico.

Tabla 3. Descripción de los conjuntos datos analizados por asignatura, año y niveles de logro académico en el estado de Tlaxcala para un total de 11 036 registros de estudiantes

Año y asignatura evaluada	Niveles de logro académico		
	0 [†]	1 [¶]	2 [§]
ESP2008 ^p	1274	5949	3813
ESP2011 [¶]	4194	5362	1480
MAT2008 ^{††}	1544	6039	3453
MAT2011 ^{¶¶}	6254	3608	1174

[†]0 = Insuficiente, [¶]1 = Elemental, [§]2 = Bueno o excelente,

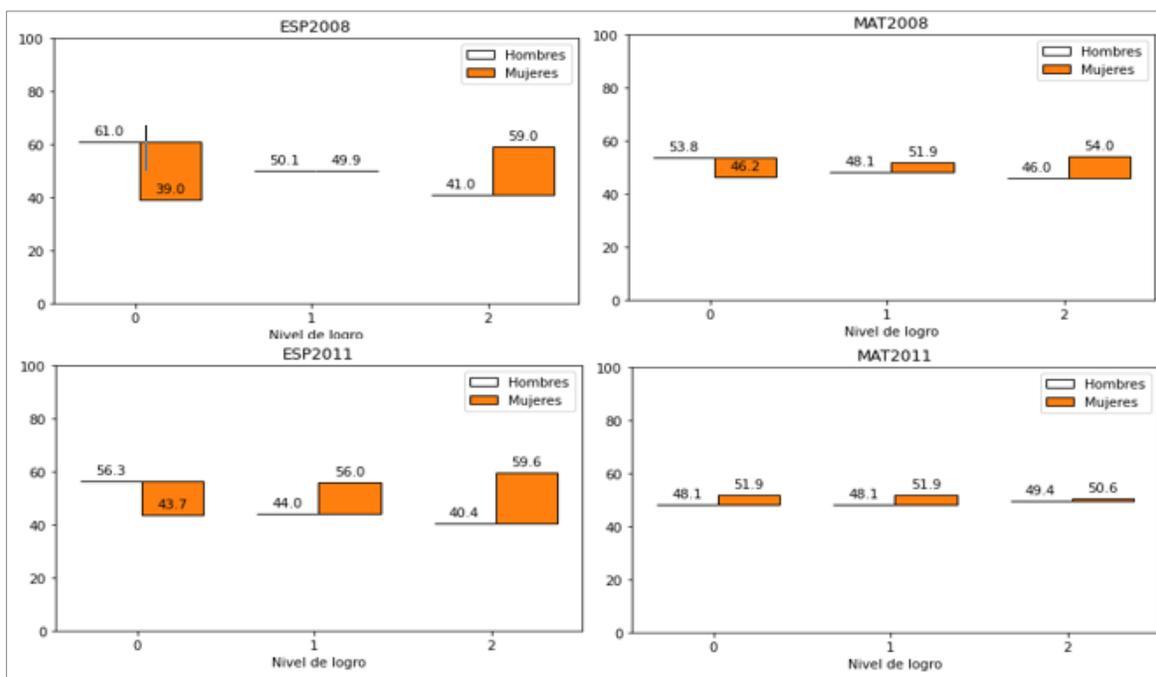
^pEspañol 2008, [¶]Español 2011, ^{††}Matemáticas 2008,

^{¶¶}Matemáticas 2011.

Fuente: Elaboración propia con base en los datos en Enlace 2008 y 2011 (SEP, 2008)

En la figura 1 se muestra la distribución de los niveles de logro por sexo. En ESP2008 hay una proporción mayor de mujeres en la clase 2 y en MAT2008 también sobresalen en las clases 1 y 2. Lo mismo se observa en ESP2011 y en MAT2011, donde las mujeres sobresalen en ambas clases.

Figura 1. Niveles de logro académico por sexo para los cuatro conjuntos de datos:
ESP2008, ESP2011, MAT2008 y MAT2011



Fuente: Elaboración propia

De la información preliminar de la prueba Enlace para Tlaxcala, se eliminaron 16 registros que tenían la misma información en todas las variables. Las variables predictoras categóricas t_esc , t_loc y n_mar se transformaron para su análisis en variables indicadoras. Los cuatro conjuntos de datos depurados consisten en 22 variables predictoras y 11 036 registros de alumnos, que representa 44.3 % de un total de 24 875 registros de alumnos evaluados en 2008, en Tlaxcala. Para el análisis, se generaron cuatro conjuntos de datos: ESP2008, ESP2011, MAT2008 y MAT2011; en cada conjunto de datos, las etiquetas de clase corresponden al nivel de logro de los alumnos obtenido en la materia correspondiente: n_esp ó n_mat .

Para el procesamiento y creación de los conjuntos de datos se usó el *software* Statistical Analysis System (SAS) v. 9.4. Para correr los algoritmos de aprendizaje automático se utilizó el *software* Python 3.8 y la biblioteca de funciones Scikit-learn para la ejecución de los códigos y la generación de resultados.

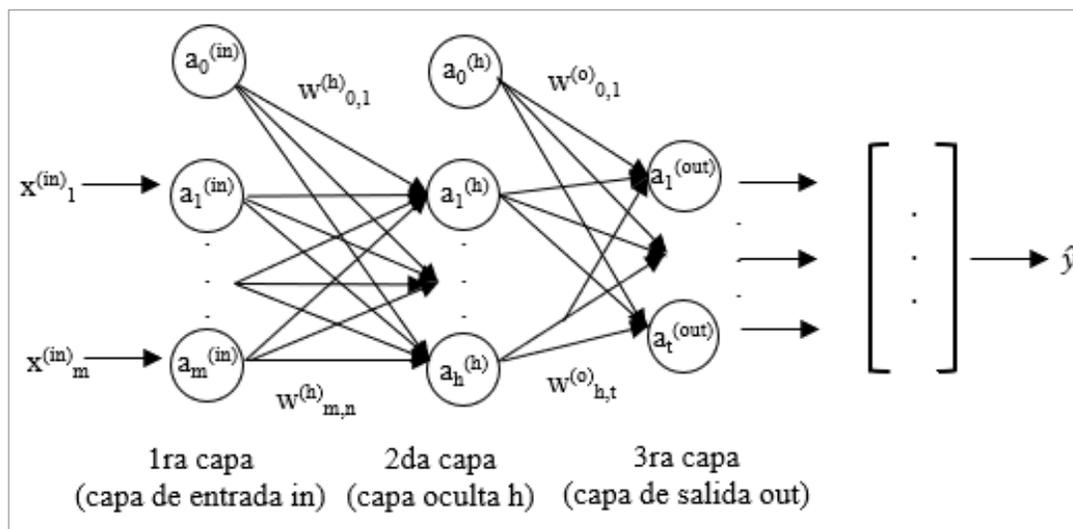
Clasificadores de aprendizaje automático

El propósito de los clasificadores de aprendizaje automático supervisado es predecir una clase objetivo a partir de variables o características de entrada. En este trabajo se implementaron tres clasificadores de aprendizaje supervisado perceptrón multicapa [MLP, por sus siglas en inglés] y potenciación del gradiente [GB, por sus siglas en inglés] para predecir el nivel del logro académico, y bosque aleatorio [RF, por sus siglas en inglés] para determinar la importancia relativa de las variables predictoras.

Perceptrón multicapa

El clasificador MLP es una red de neuronas conectadas por pesos o parámetros, estructuradas en una capa de entrada (*in*), una o más capas ocultas (*h*) y una capa de salida (*out*). La arquitectura básica de MLP consiste en tres capas (figura 2). Entre más capas tiene la red, esta es más compleja, y adquiere mayor habilidad para resolver problemas complejos (Borkar y Rajeswari, 2014).

Figura 2. Arquitectura del clasificador perceptrón multicapa



Fuente: Raschka y Mirjalili (2017)

El funcionamiento del clasificador MLP consiste en, que dado un vector de datos de entrada x_i con m variables predictoras, en la capa oculta se infieren M funciones; luego, en la capa de salida se determina la respuesta predicha, al aplicar las funciones inferidas en la capa oculta por medio de una transformación no lineal (González *et al.*, 2012). Todas las capas se conectan hacia adelante, y se representan por una matriz de pesos W , que se inicializa

con valores aleatorios pequeños. Primero se activa la capa oculta $a_1^{(h)}$ por medio de una función de activación ϕ (por ejemplo, una función sigmoide) a los valores de Z (matriz de valores netos), que resulta de la combinación lineal de las variables de entrada con los pesos W ; estos valores son las entradas de la capa de salida. La activación de la capa oculta se realiza con las siguientes expresiones:

$$Z^{(h)} = A^{(in)}W^{(h)}, A^{(h)} = \phi(Z^{(h)})$$

donde $A^{(in)}$ es una matriz de características o muestras $x^{(in)}$; $W^{(h)}$ es la matriz de pesos, y $\phi(\cdot)$ es la función de activación. De manera similar, se genera la activación de la capa de salida:

$$Z^{(out)} = A^{(h)}W^{(out)}, A^{(out)} = \phi(Z^{(out)})$$

donde $W^{(out)}$ es una matriz de pesos de salida; y $A^{(out)}$ es una matriz de probabilidades con las respuestas o clases predichas de la red. Para determinar el error de clasificación, se compara la clase objetivo con la clase predicha. El algoritmo de propagación hacia atrás se utiliza para distribuir los errores y se obtienen las derivadas parciales con respecto a los pesos de la red para actualizar el modelo (Raschka y Mirjalili, 2017).

Clasificador potenciación del gradiente

El clasificador GB consiste en un conjunto de árboles de decisión individuales que se entrenan de manera secuencial, de tal forma que cada árbol mejore los errores de los árboles anteriores. Para predecir una nueva observación se agregan las predicciones de todos los árboles individuales del modelo. GB puede emplear cualquier función de pérdida siempre que ésta, sea diferenciable. Se ajusta un modelo, por ejemplo, f_1 para predecir la variable respuesta, después se calculan los errores $y - f_1(x)$; luego, se ajusta un modelo f_2 que intenta predecir los errores del modelo anterior; nuevamente se ajusta un modelo f_3 que trata de corregir los errores de los modelos anteriores y esto se repite m veces. Para evitar el sobreajuste del modelo, se emplea un parámetro de regularización, que se denomina *tasa de aprendizaje* (λ), que limita la influencia de cada modelo en el conjunto de ensamble.

$$y \approx \lambda \sum_{i=1}^m f_i(x)$$

La idea que está detrás de la potenciación es ajustar de forma secuencial múltiples modelos sencillos, en donde cada modelo emplea información del modelo anterior para “aprender de sus errores” y mejorar en cada iteración; como valor final se toma el promedio de las predicciones (Rogers y Gunn, 2005).

Bosque aleatorio

RF es un conjunto de árboles de decisión y su objetivo es promediar múltiples árboles de decisión para construir un modelo más robusto que tenga una mejor generalización y sea menos susceptible al sobreajuste (Raschka y Mirjalili, 2017). Para predecir la clase se usan las reglas de cada árbol y se asignan por mayoría de votos (Breiman, 2001). El algoritmo RF se resume como sigue:

Se selecciona una muestra de tamaño n del conjunto de variables predictoras (mediante muestreo aleatorio sin reemplazo, *bootstrap*), el árbol crece a partir de una muestra inicial; para cada nodo se seleccionan d características sin reemplazo; se divide el nodo con la función que proporcione la mejor división de acuerdo con la función objetivo ganancia de información (IG, por sus siglas en inglés), que se define por:

$$IG = (D_p, f) = I(D_p) - \frac{N_l}{N_p} I(D_l) - \frac{N_r}{N_p} I(D_r)$$

donde f es la característica para realizar la división; N_p es el número total de muestras en el nodo padre; N_l es el número de muestras en el nodo hijo izquierdo; N_r es el número de muestras en el nodo hijo derecho; I es la medida de impureza (gini, entropía o error de clasificación); D_p es el conjunto de datos en el nodo padre; D_l es el conjunto de datos en el nodo hijo izquierdo, y D_r es el conjunto de datos en el nodo hijo derecho.

RF también se usa para determinar la importancia de un conjunto de variables en el modelo. El algoritmo RF crea clasificadores con una selección aleatoria de las características; esto logra una buena exploración de subconjuntos de éstas, en donde se seleccionan aquellas variables con mayor importancia (Rogers y Gunn, 2005).

Criterios de desempeño de los modelos de predicción

Para evaluar el desempeño de los modelos de clasificación MLP, GB y RF las métricas se obtienen a partir de una matriz de confusión (MC) la cual, describe el conteo de los verdaderos positivos (VP), verdaderos negativos (VN), falsos positivos (FP) y falsos negativos (FN). Las filas representan el número de muestras en la clase observada y las columnas el número de predicciones de cada clase. La diagonal de MC corresponde al número de muestras que el algoritmo clasifica correctamente en cada clase. Si MC solo tiene valores positivos en la diagonal, indica que el clasificador clasifica correctamente todas las muestras. La métrica precisión global de clasificación (PG) mide la proporción global de muestras bien clasificadas en cada clase y se calcula como:

$$PG = \frac{VP + VN}{FP + FN + VP + VN}$$

Las métricas para medir el desempeño del clasificador en cada clase son precisión (P), sensibilidad (S), especificidad (E) y puntaje $F1$. Se definen con las siguientes expresiones:

$$P = \frac{VP}{VP + FP}$$

$$S = \frac{VP}{FN + VP}$$

$$E = \frac{VN}{VN + FP}$$

$$F1 = 2 \frac{P \times S}{P + S}$$

En este caso, el valor de $F1$ resume P y S en una sola métrica, es un estimador apropiado en clases desbalanceadas y varía entre cero y uno. La curva *receiver operating characteristics* (ROC) es una curva que relaciona valores de S versus $1-E$. Los diferentes puntos en la curva corresponden a los puntos de corte utilizados para determinar si los resultados de la prueba son positivos. El valor de AUC_{ROC} (área bajo la curva ROC) se interpreta como la probabilidad de que en dos muestras, una positiva y una negativa, la prueba asigne una probabilidad más alta a la muestra positiva, clasificación correcta (Mandrekar, 2010). Su valor oscila entre cero y uno; cuanto mayor es AUC_{ROC} mejor es la clasificación, un valor cercano a 0.50 indica una mala clasificación. La curva $P-S$ es el resultado de graficar P versus S . Ésta permite observar a partir de qué S se tiene una degradación de P y viceversa. El resultado ideal es una curva que se acerque a la esquina superior derecha (alta P y S), lo que

genera un área bajo la curva AUC_{P-S} , que entre más cercano a uno, es mejor el modelo (Saito y Rehmsmeier, 2015).

Entrenamiento y validación de los modelos

Para realizar el entrenamiento de los modelos, cada conjunto de datos se dividió en dos particiones aleatorias, 80 % para entrenamiento y 20 % prueba. Para cada conjunto de datos (ESP2008, MAT2008, ESP2011 y MAT2011) se aplicó el clasificador RF para evaluar la importancia relativa de las características de entrada. En la prueba de los modelos se consideró 20 % de los datos y se presenta el caso de clases desbalanceadas en los extremos (tabla 4).

Tabla 4. Número de observaciones de los conjuntos de prueba por clase objetivo o nivel de logro académico, en español y matemáticas 2008 y 2011

Clase	ESP2008	ESP2011	MAT2008	MAT2011
0 [†]	255	839	309	1251
1 [¶]	1190	1073	1208	722
2 [§]	763	296	691	235

[†]0: insuficiente; [¶]1: elemental; [§]2: bueno o excelente.

Fuente: Elaboración propia, con base en datos de SEP (2008)

Selección de hiperparámetros óptimos

En los modelos de aprendizaje automático, los hiperparámetros óptimos se ajustan en el entrenamiento para obtener el mejor desempeño. La selección de los hiperparámetros óptimos consiste en encontrar la combinación de valores de los hiperparámetros que maximice el desempeño del clasificador con base en una métrica, en este estudio se utilizó PG. La selección de hiperparámetros para cada clasificador se realizó por medio de una validación cruzada (CV). El entrenamiento se realizó con una muestra aleatoria de 80 % del conjunto total de datos. El método CV consiste en subdividir de forma aleatoria el conjunto de entrenamiento en k subconjuntos disjuntos del mismo tamaño. Luego, para cada combinación de valores de los hiperparámetros (tabla 5), el modelo se ejecuta k veces. En cada iteración k se usa uno de los subconjuntos disjuntos como conjunto de validación y los restantes como conjunto de entrenamiento (80 % entrenamiento y 20 % validación) y se

obtiene un valor de la métrica de desempeño PG. Después de evaluar diferentes combinaciones de valores de los hiperparámetros, se selecciona la combinación de valores de los hiperparámetros que maximice la PG promedio obtenida de VC ($k = 5$).

Tabla 5. Valores para la búsqueda y selección de hiperparámetros de los clasificadores perceptrón multicapa (MLP) y potenciación del gradiente (GB)

Modelo	Hiperparámetros	Valores considerados
MLP	<i>nco</i>	50, 100, 200, 250
	<i>fa</i>	<i>tanh, relu, logistic</i>
	<i>op</i>	<i>sgd, adam, lbfgs</i>
	<i>re</i>	0.0001, 0.001, 0.1
	<i>ta</i>	<i>constant, adaptative</i>
	<i>mi</i>	200
GB	<i>mi</i>	90, 100, 110
	<i>pa</i>	4, 5, 6
	<i>ha</i>	3, 4

nco: neuronas en capa oculta; *fa*: función de activación; *op*: optimizador de pesos; *re*: regularizador; *ta*: tasa de aprendizaje; *mi*: máximo de iteraciones; *pa*: profundidad de árbol; *ha*: hojas por árbol.

Fuente: Elaboración propia

Posterior a la etapa de selección de los hiperparámetros óptimos de los clasificadores MLP y GB, se realizó la evaluación final del desempeño de éstos. Para determinar la PG de cada modelo se consideró el conjunto de datos completo de cada escenario de entrada (tabla 3), la combinación óptima de hiperparámetros seleccionada y el procedimiento VC ($k = 5$). En cada iteración se obtiene un valor PG del modelo; al finalizar las k iteraciones, se calculó el promedio de los valores de PG y su desviación estándar, y las demás métricas propuestas en el estudio.

Importancia relativa de las características de entrada

RF se utilizó para determinar la importancia relativa de las características o variables de entrada para predecir la clase objetivo. RF construye una gran cantidad de clasificadores con base en subconjuntos de variables que se seleccionan al azar. En cada nodo de RF se selecciona una variable de entrada que se usa para dividir el nodo y maximizar la ganancia de información (métrica de desempeño). Las medidas de importancia de variables se utilizan para determinar el desempeño del modelo de aprendizaje automático (Rogers y Gunn, 2005).

Para calcular la importancia relativa de las 13 variables predictoras con el modelo RF, se aplicó el procedimiento VC ($k = 5$) para seleccionar los hiperparámetros óptimos. Posteriormente, se aplicó la opción *feature importance* al modelo RF por medio de la librería Scikit-learn de Python. Esta etapa se realizó para los cuatro conjuntos de datos: ESP2008, ESP2011, MAT2008 y MAT2011.

Resultados

Hiperparámetros óptimos

Los hiperparámetros óptimos seleccionados con validación cruzada y búsqueda en retícula de los clasificadores MLP y GB para cada conjunto de datos se ilustran en la tabla 6. Los hiperparámetros óptimos de cada clasificador, en general, dependen del conjunto de datos analizado.

Tabla 6. Hiperparámetros óptimos de los clasificadores perceptrón multicapa (MLP) y potenciación del gradiente (GB) para cada escenario de entrada de datos

Modelo	Hiperparámetros	ESP2008	ESP2011	MAT2008	MAT2011
MLP	<i>Nco</i>	100	200	50	30
	<i>fa</i>	<i>logistic</i>	<i>tanh</i>	<i>relu</i>	<i>logistic</i>
	<i>op</i>	<i>adam</i>	<i>sgd</i>	<i>sgd</i>	<i>adam</i>
	<i>re</i>	0.1	0.1	0.1	0.0001
	<i>ta</i>	<i>constant</i>	<i>adaptative</i>	<i>constant</i>	<i>constant</i>
GB	<i>mi</i>	90	90	90	110
	<i>pa</i>	4	4	4	4
	<i>ha</i>	3	4	3	4

nco: neuronas en capa oculta; *fa*: función de activación; *op*: optimizador de pesos; *re*: regularizador; *ta*: tasa de aprendizaje; *mi*: máximo de iteraciones; *pa*: profundidad de árbol; *ha*: hojas por árbol.

Fuente: Elaboración propia

Desempeño de los clasificadores

El desempeño promedio y su desviación estándar en predicción, de los modelos MLP y GB para los cuatro escenarios analizados de la prueba Enlace, en términos de la precisión global de clasificación MLP fue superior a GB con el escenario ESP2008 y GB fue superior a MLP con el escenario MAT2008 (tabla 7).

Tabla 7. Precisión global (*PG*) en predicción promedio (+/- desviación estándar) de los clasificadores perceptrón multicapa (MLP) y potenciación del gradiente (GB) en español y matemáticas

Modelo	ESP2008	ESP2011	MAT2008	MAT2011
MLP	0.701(+/- 0.033)	0.611(+/- 0.019)	0.672(+/- 0.034)	0.631(+/- 0.008)
GB	0.695(+/- 0.024)	0.610(+/- 0.016)	0.688(+/- 0.008)	0.635(+/- 0.006)

Fuente: Elaboración propia

El clasificador MLP obtuvo mejor desempeño global que GB para predecir el grado de logro académico de los alumnos en español 2008 y 2011, con una PG de 70.1 % y 61.1 %

respectivamente; sin embargo, el desempeño de ambos clasificadores es muy similar en ambas asignaturas en 2008 y 2011 (tabla 7).

En la asignatura de español, MLP y GB obtuvieron mejor desempeño en 2008 para clasificar las clases 1 y 2, que la clase 0 (*FI* de 77.0 % y 75.0% *versus* 33.0 %, respectivamente). Por otro lado, en 2011, MLP y GB obtuvieron mejor desempeño para clasificar las clases 0 y 1, que la clase 2. Lo anterior se refuerza al observar los resultados de AUC_{P-S} con valores de 0.70 y 0.60 para las clases 0 y 1 (tabla 8, figura 3).

Tabla 8. Desempeño en predicción de los clasificadores perceptrón multicapa (MLP) y potenciación del gradiente (GB) por clase de logro académico en español 2008 y 2011

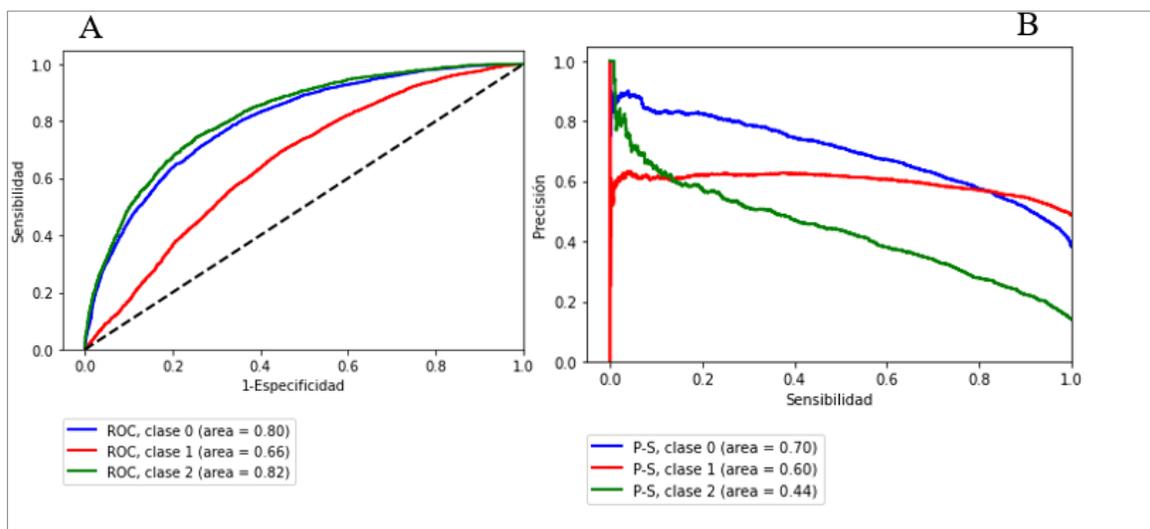
Modelo	Clase	ESP2008			ESP2011		
		<i>FI</i>	AUC_{ROC}	AUC_{P-S}	<i>FI</i>	AUC_{ROC}	AUC_{P-S}
MLP	0	0.32	0.86	0.46	0.65	0.80	0.70
	1	0.77	0.74	0.72	0.67	0.66	0.60
	2	0.75	0.87	0.75	0.41	0.82	0.44
GB	0	0.33	0.86	0.40	0.65	0.80	0.69
	1	0.77	0.71	0.68	0.67	0.64	0.57
	2	0.75	0.87	0.77	0.42	0.81	0.41

FI: f1-score, AUC_{ROC} : área bajo la curva ROC, AUC_{P-S} : área bajo la curva *P-S*, clases 0: insuficiente, 1: elemental, 2: bueno o excelente

Fuente: Elaboración propia

MLP y GB obtuvieron bajo desempeño para clasificar la clase 2 con $FI = 0.41$. Esto se confirma al observar las curvas *P-S* con un valor bajo de AUC_{P-S} para la clase 2 (figura 3).

Figura 3. Curvas de desempeño en predicción del clasificador perceptrón multicapa (MLP) en español 2011 por clase (0: insuficiente, 1: elemental, 2: bueno o excelente) A: curvas ROC, B: curvas P-S



Fuente: Elaboración propia

Para el conjunto de datos de matemáticas 2008, GB obtuvo mejor desempeño que MLP, con una PG promedio de 68.8 % (tabla 3). Las clases 1 y 2 ($F1 = 77.0\%$ y 74.0%) resultan mejor clasificadas que la clase 0 ($F1 = 43\%$). MLP obtuvo AUC_{P-S} de 0.68 y 0.73 para las clases 1 y 2, respectivamente, y 0.50 para la clase 0. GB obtuvo un desempeño muy próximo al MLP.

Con el conjunto de datos de matemáticas de 2011, GB fue ligeramente superior a MLP (GB obtuvo una PG promedio de 63.5 %). GB tuvo mejor desempeño que MLP para clasificar la clase 0 con $F1$ de 77 %. Ambos modelos tuvieron bajo desempeño para clasificar las clases 1 y 2. De los resultados de AUC_{P-S} , se constata que ambos modelos tuvieron buen desempeño para clasificar la clase 0, y bajo para las clases 1 y 2 (0.37) (tabla 9, figura 4).

Tabla 9. Desempeño en predicción de los clasificadores perceptrón multicapa (MLP) y potenciación del gradiente (GB) por clase de logro académico en matemáticas 2008 y 2011

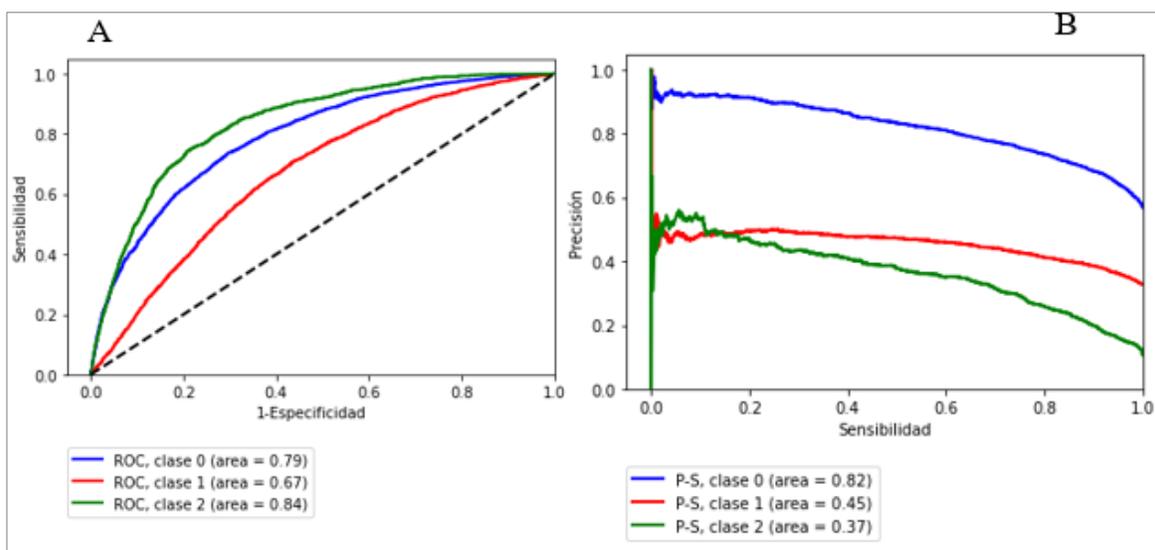
Modelo	Clase	MAT2008			MAT2011		
		<i>FI</i>	AUC_{ROC}	AUC_{P-S}	<i>FI</i>	AUC_{ROC}	AUC_{P-S}
MLP	0	0.43	0.86	0.50	0.76	0.79	0.82
	1	0.77	0.70	0.68	0.44	0.68	0.47
	2	0.74	0.86	0.73	0.30	0.83	0.38
GB	0	0.39	0.86	0.48	0.77	0.79	0.82
	1	0.78	0.70	0.69	0.46	0.67	0.45
	2	0.73	0.87	0.76	0.39	0.84	0.35

P: precisión; *FI*: f1-score, AUC_{ROC} : área bajo la curva ROC, AUC_{P-S} : área bajo la curva *P-S*, clases 0: insuficiente, 1: elemental, 2: bueno o excelente

Fuente: Elaboración propia

Figura 4. Curvas de desempeño en predicción del clasificador potenciación del gradiente (GB) en matemáticas 2011 por clase (0: insuficiente, 1: elemental, 2: bueno o excelente),

A: curvas ROC, B: curvas *P-S*

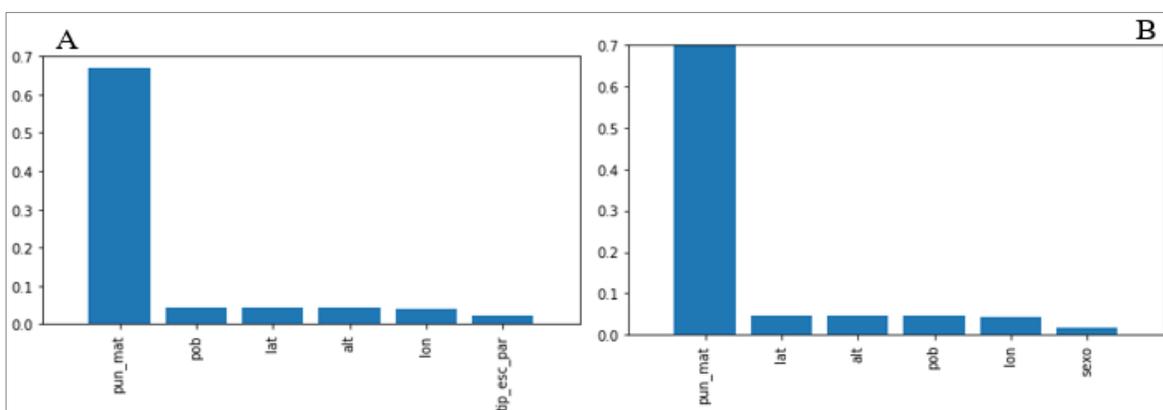


Fuente: Elaboración propia

Importancia relativa de características

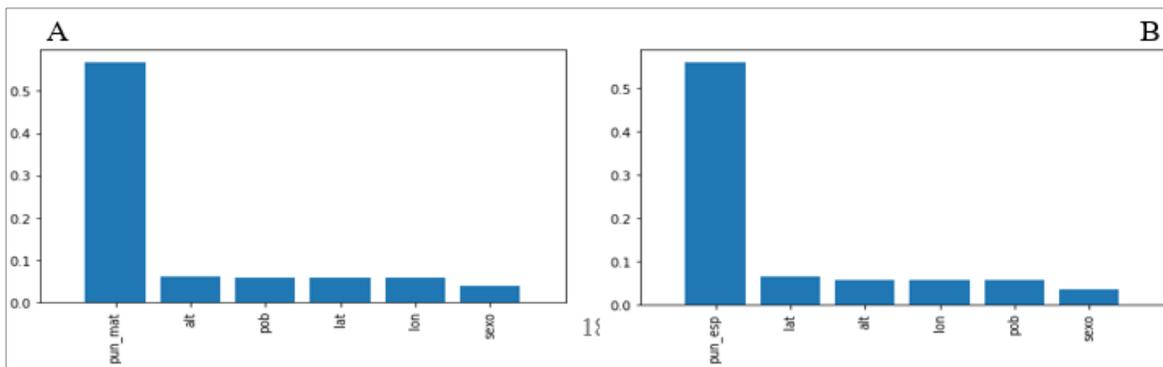
Entre los resultados de importancia relativa de las variables de entrada, determinados con el clasificador RF para los cuatro conjuntos de datos (español y matemáticas, 2008 y 2011), se encontró que las variables de puntaje *pun_esp* y *pun_mat* son las más importantes para predecir el logro académico (figuras 5 y 6). Esto es, existe una fuerte asociación (0.65) entre el puntaje obtenido por un alumno en matemáticas y su grado de logro académico en español (véase los incisos *a*) de las figuras 5 y 6). Similarmente, existe una fuerte asociación (0.7) entre el puntaje obtenido por un alumno en español y el grado de logro académico en matemáticas (véase los incisos *b*) de las figuras 5 y 6). Esto se preserva al inicio (2008) y al término del periodo analizado (2011).

Figura 5. Importancia relativa de variables de entrada obtenida con el clasificador bosque aleatorio (RF) A: español 2008, B: matemáticas 2008



Fuente: Elaboración propia

Figura 6. Importancia relativa de variables de entrada, obtenida con el clasificador bosque aleatorio (RF) A: español 2011, B: matemáticas 2011



Fuente: Elaboración propia

Además de las variables que corresponden a los puntajes en la prueba, las variables de *sexo*, *t_esc* particular, *pob*, *lat*, *lon* y *alt* resultan importantes para los modelos mediante RF. Esto es, son variables que influyen en la clasificación de los alumnos en alguno de los niveles de logro analizados.

Discusión

Los resultados obtenidos muestran la asociación de las 13 variables de entrada con las tres clases o niveles de logro académico (insuficiente, elemental, y bueno o excelente) obtenidos por los alumnos del estado de Tlaxcala. Una de las variables más importantes fue el puntaje obtenido en matemáticas para predecir el grado o nivel de logro académico en español y el puntaje obtenido en español para predecir el grado de logro académico en matemáticas. En orden de importancia, aunque en menor medida, también figuran la ubicación geográfica de la escuela, la población de la localidad de ubicación de la escuela y el sexo del alumno. Se destaca la influencia que se observó de las variables *pun_esp* y *pun_mat* en los resultados obtenidos para la clasificación de los alumnos en los niveles de logro.

Fernández (2003) señala que el aprendizaje del idioma español resulta de un proceso de acumulación de experiencias pedagógicas que el alumno tiene en su estancia en la escuela y al aprendizaje de matemáticas como un proceso constructivo que se relaciona con la formulación o comprensión de conceptos con la resolución de problemas. Es importante observar las calificaciones de los estudiantes en español para poder predecir la clasificación

de los niveles de logro en matemáticas y viceversa, es decir, conviene saber que pasó en la otra materia como una medida general de la capacidad de los alumnos.

Con las diferentes métricas seleccionadas para comparar los dos modelos, el área bajo la curva AUC_{P-S} proporcionó más información para evaluar el desempeño de los clasificadores al discriminar el porcentaje de muestras clasificadas correctamente y los resultados se reflejan similarmente en la matriz de confusión.

Los dos algoritmos de aprendizaje automático (MLP y GB), obtuvieron un desempeño global de clasificación correcta (PG) mayor a 60.0 %). Una limitación del trabajo para incrementar el desempeño de los clasificadores fue la presencia de clases objetivo desbalanceadas, el clasificador tiende a dar mayor importancia a las clases mayoritarias. Para mejorar el trabajo, se pueden considerar variables de contexto adicionales y observar si éstas mejoran la clasificación. Álvarez *et al.* (2007) utilizaron variables asociadas a los estudiantes referentes a indicadores socioeconómicos, características de la escuela y aspectos institucionales (pedagogía estatal, influencia sindical, etcétera) para determinar qué factores influyen en el desempeño escolar de matemáticas, ciencias y lectura de PISA, igualmente, Hussain y Qasim (2021) utilizaron datos históricos de calificaciones para predecir las calificaciones de los estudiantes por medio de algoritmos de aprendizaje automático.

Conclusiones

Los clasificadores de aprendizaje automático perceptrón multicapa (MLP) y potenciación del gradiente (GB) obtuvieron desempeños comparables en términos de la precisión global de clasificación (PG) para predecir los niveles de logro académico (0: insuficiente, 1: elemental, y 2: bueno o excelente) de estudiantes de nivel básico y medio del estado de Tlaxcala, a partir de variables contextuales extraídas de la prueba Enlace (Evaluación Nacional de Logro Académico en Centros Escolares). En matemáticas, GB obtuvo una PG de 68.8 % en 2008, y 63.5 % en 2011; asimismo, en 2008, MLP y GB obtuvieron mejor desempeño para clasificar las clases 1 y 2 que la clase 0 (insuficiente). En cambio, en 2011, en ambas asignaturas, MLP y GB obtuvieron mejor desempeño para clasificar las clases 0 y 1 que la clase 2.

Las variables contextuales utilizadas en este estudio mostraron una asociación con los niveles de logro académico; en particular, las variables becario, sexo y turno de la escuela. El puntaje en español obtenido por un estudiante influye en el nivel de logro académico en matemáticas y viceversa. Estos resultados muestran la importancia de los algoritmos de aprendizaje automático para identificar factores relevantes que inciden en el rendimiento escolar de los estudiantes a partir del análisis de datos masivos de información escolar existente en la Secretaría de Educación Pública.

Los clasificadores GB y MLP representan un enfoque alternativo para identificar las variables o factores contextuales que favorecen o limitan el logro académico de los estudiantes y constituyen una herramienta de apoyo a la toma de decisiones para identificar alumnos con bajo desempeño y plantear soluciones focalizadas a problemas estructurales como la deserción escolar.

Futuras líneas de investigación

Para complementar el trabajo es importante considerar otras variables contextuales como la escolaridad de los padres, características de la familia, características de la escuela, etcétera, y que puede obtenerse por medio del cruce de información de rendimiento escolar de la prueba Enlace con los cuestionarios de contexto que se levantaron a la par de la prueba.

Estas nuevas variables pueden ser analizadas con los modelos aprendizaje automático para evaluar su influencia en el rendimiento escolar. Además, es posible probar otros algoritmos de clasificación como las máquinas de soporte vectorial (SVM, por sus siglas en inglés) o el k-vecino más cercano (KNN, por sus siglas en inglés). Igualmente, con el enfoque aplicado en este estudio, es posible seleccionar información escolar de otras entidades o regiones del país para evaluar el rendimiento escolar en otros contextos socioeconómicos.

Agradecimientos

Al Consejo Nacional de Ciencia y Tecnología (Conacyt) por el apoyo económico brindado para la realización de esta investigación.

Referencias

- Altabrawee, H., Osama, A. J. and Qaisir, A. S. (2019). Predicting Students' Performance Using Machine Learning Techniques. *Journal of University of Babylon for Pure and Applied Sciences*, 27(1), 194-205. Retrieved from <https://doi.org/10.29196/jubpas.v27i1.2108>.
- Álvarez, J., García, M. V. and Patrinos, H. A. (2007). *Institutional Effects as Determinants of Learning Outcomes. Exploring State Variations in Mexico*. Washington, United States: The World Bank.
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), 5-32. Retrieved from <https://doi.org/10.1023/A:1010933404324>.
- Borkar, S. and Rajeswari, K. (2014). Attributes Selection for Predicting Students' Academic Performance Using Education Data Mining and Artificial Neural Network. *International Journal of Computer Applications*, 86(10), 25-29. Retrieved from <https://doi.org/10.5120/15022-3310>.
- Dambić, G., Krajcar, M. and Bele, D. (2016). Machine learning model for early detection of higher education students that need additional attention in introductory programming courses. *International Journal of Digital Technology & Economy*, 1(1), 1-11.
- Fernández, T. (2003). *Determinantes sociales, organizacionales, e institucionales de los aprendizajes en la educación primaria en México: un análisis de tres niveles (2001)*. México: Instituto Nacional para la Evaluación de la Educación. Recuperado de <https://www.inee.edu.mx/wp-content/uploads/2019/01/P1C126.pdf>.
- González, J. M., de los Campos, G., Pérez, P., Gianola, D., Cairns, J. E., Mahuku, G., Babu, R. and Crossa, J. (2012). Genome-enabled prediction of genetic values using radial basis function neural networks. *Theoretical and Applied Genetics*, 125(4), 759-771. Retrieved from <https://doi.org/10.1007/s00122-012-1868-9>.
- Hussain, S., and Qasim, K. M. (2021). Predicting Students' Academic Performance at Secondary and Intermediate Level Using Machine Learning. *Annals of Data Science*. Retrieved from <https://doi.org/10.1007/s40745-021-00341-0>.
- Instituto Nacional de Estadística y Geografía [Inegi]. (2020). Catálogo Único de Claves de Áreas Geoestadísticas Estatales, Municipales y Localidades. Recuperado de <https://www.inegi.org.mx/app/ageeml/#>.

- Instituto Nacional para la Evaluación de la Educación [INEE]. (2019). Evaluaciones. Bases de datos. Recuperado de <https://www.inee.edu.mx/evaluaciones/bases-de-datos/>.
- Mandrekar, J. (2010). Receiver Operating Characteristic Curve in Diagnostic Test Assessment. *Journal of Thoracic Oncology*, 5(9), 1315-1316. Retrieved from <https://doi.org/10.1097/JTO.0b013e3181ec173d>.
- Martínez, F. (coord.) (2015). *Las pruebas Enlace y Excale. Un estudio de validación*. Ciudad de México, México: Instituto Nacional para la Evaluación de la Educación.
- Organisation for Economic Co-operation and Development [OCDE]. (2005). School Factors Related to Quality and Equity: Results from PISA 2000. Paris, France: Organisation for Economic Co-operation and Development. Retrieved from https://read.oecd-ilibrary.org/education/school-factors-related-to-quality-and-equity_9789264008199-en#page1.
- Rai, S., Shastry, K., Pratap, S., Kishore, S., Mishara, P. and Sanjaty, H. (2020). Machine Learning Approach for Student Academic Performance Prediction. In Bhateja, V., Grobelnik, M., Peng, S., Sataphaty, S. and Zhang, Y. (eds.), *Evolution in Computational Intelligence. Frontiers in Intelligent Computing: Theory and Applications* (pp. 611-618). Singapore: Springer.
- Raschka, S. and Mirjalili, V. (2017). *Python Machine Learning. Machine Learning and Deep Learning with Python-scikit-learn, and TensorFlow*. Birmingham, England: Packt Publishing.
- Rogers, J. and Gunn, S. (2005). Identifying Feature Relevance Using a Random Forest. In Saunders, C., Grobelnik, M., Gunn, S. and Shawe-Taylor, J. (eds.), *Subspace, Latent, Structure and Feature Selection* (pp. 173-184). Bohinj, Slovenia: Springer.
- Saito, T. and Rehmsmeier, M. (2015). The Precision-Recall Plot is More Informative than the ROC Plot when Evaluating Binary Classifiers on Imbalanced Datasets. *PLoS ONE*, 10(3), 1-21. Retrieved from <https://doi.org/10.1371/journal.pone.0118432>.
- Secretaría de Educación Pública [SEP]. (2008). *Enlace. Educación básica. Manual técnico 2008*. México: Secretaría de Educación Pública. Recuperado de http://enlace.sep.gob.mx/ba/manuales_tecnicos/.SEP-IEIA.

Rol de Contribución	Autor (es)
Conceptualización	Juan Manuel González-Camacho (principal), Miguel Angel Morales Hernández (que apoya)
Metodología	Miguel Ángel Morales (igual) -Hernández, Juan Manuel González-Camacho (igual)
Software	Miguel Ángel Morales -Hernández (principal), Juan Manuel González-Camacho (que apoya)
Validación	Miguel Ángel Morales -Hernández (principal), Juan Manuel González-Camacho (que apoya)
Análisis Formal	Miguel Ángel Morales -Hernández (igual), Juan Manuel González-Camacho (igual)
Investigación	Miguel Ángel Morales -Hernández (principal), Juan Manuel González Camacho (que apoya)
Recursos	Miguel Ángel Morales -Hernández (igual), Juan Manuel González-Camacho (igual)
Curación de datos	Miguel Ángel Morales Hernández
Escritura - Preparación del borrador original	Miguel Ángel Morales -Hernández (igual), Juan Manuel González-Camacho (igual)
Escritura - Revisión y edición	Miguel Ángel Morales -Hernández (igual), Juan Manuel González-Camacho (igual), David H. Del Valle Paniagua (que apoya), Héctor Robles (que apoya), José Rafael Durán Moreno (que apoya)
Visualización	Miguel Angel Morales Hernández
Supervisión	Juan Manuel González-Camacho
Administración de Proyectos	Juan Manuel González-Camacho
Adquisición de fondos	Miguel Ángel Morales -Hernández (igual), Juan Manuel González-Camacho (igual)