

ARTÍCULO DE INVESTIGACIÓN

Acceso abierto



Predicción de estudiantes en riesgo de fracaso académico utilizando un modelo de conjunto durante una pandemia en un sistema de aprendizaje a distancia

Halit Karalar^{1*} , Ceyhun Kapucu² Hüseyn Gürüler³

*Correspondencia:
halit@mu.edu.tr

¹ Departamento de Instrucción y Educación Informática
 Tecnologías, Facultad de Educación, Muğla Sıtkı
 Universidad de Kocman, Muğla,
 pavo
 La lista completa de la información
 del autor está disponible al final del
 artículo

Abstracto

Predecir los estudiantes en riesgo de fracaso académico es valioso para que las instituciones de educación superior mejoren el desempeño de los estudiantes. Durante la pandemia, con la transición a la educación a distancia obligatoria en la educación superior, se ha vuelto aún más importante identificar a estos estudiantes y realizar intervenciones educativas para evitar dejarlos atrás. Este objetivo se puede lograr mediante nuevas técnicas de minería de datos y métodos de aprendizaje automático. Este estudio tomó en cuenta las características de actividad sincrónica y asincrónica de los estudiantes para identificar a los estudiantes en riesgo de fracaso académico durante la pandemia. Además, este estudio propone un modelo de conjunto óptimo que predice a los estudiantes en riesgo utilizando una combinación de algoritmos de aprendizaje automático relevantes. Se predijeron las actuaciones de más de dos mil estudiantes universitarios con un modelo de conjunto en términos de género, grado, número de notas de lectura descargadas y materiales del curso, tiempo total dedicado a las sesiones en línea, número de asistencias y puntuación de la prueba. Las actividades de aprendizaje asincrónico resultaron más determinantes que las sincrónicas. El modelo de conjunto propuesto hizo una buena predicción con una especificidad del 90,34 %. Por lo tanto, se sugiere a los profesionales que supervisen y organicen las actividades de formación en consecuencia.

Palabras clave: Predicción del desempeño de los estudiantes, Predicción de estudiantes en riesgo, Modelo de aprendizaje conjunto, Minería de datos educativos, Aprendizaje a distancia, Pandemia de COVID-19, Educación en pandemia

Introducción

Después de que la Organización Mundial de la Salud declarara el coronavirus COVID-19 como una pandemia mundial el 11 de marzo de 2020 (OMS, 2020), los gobiernos tuvieron que tomar decisiones estratégicas para hacer frente al virus. Casi todos los campos de la vida se vieron afectados por estas decisiones. Las instituciones educativas se cerraron temporalmente en muchos países. Según datos de la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura del 2 de abril de 2020 (UNESCO, 2020), aproximadamente 1500 millones de estudiantes (alrededor del 85 % a nivel mundial) en 172 países se vieron afectados por el cierre de las escuelas. En Turquía, con decisiones tomadas por el Consejo de Educación Superior (2020), la educación a distancia obligatoria (CDE) comenzó en las universidades el 13 de marzo de 2020. Esta decisión aún es válida en la primera mitad de 2021.

Para las instituciones de educación superior, predecir los estudiantes en riesgo durante la pandemia se vuelve más importante ya que los estudiantes pueden sentirse aislados durante la CDE. Para reducir este riesgo, es importante dar a los estudiantes retroalimentación alentadora y de apoyo a tiempo. Los estudiantes en riesgo deben anticiparse primero durante CDE. Durante más de diez años, los investigadores están tratando de desarrollar una solución mediante técnicas de minería de datos (DM) y aprendizaje automático (ML) que pueda analizar y predecir el rendimiento de los estudiantes y su causa raíz (Injadat et al., 2020; Romero & Ventura, 2013). Hasta donde sabemos, no existe ningún estudio que prediga a los estudiantes en riesgo de fracaso académico durante la pandemia.

Para mejorar el desempeño de los estudiantes y la eficiencia y eficacia de la educación superior, la predicción de los estudiantes en riesgo de fracaso académico también es esencial para las intervenciones educativas oportunas (Adejo & Connolly, 2017; Helal et al., 2018). Los estudios para predecir el desempeño de los estudiantes han cobrado un impulso significativo en la última década (Abu Saa et al., 2019). Estas predicciones se centran principalmente en modelos de clasificación y regresión. La clasificación (aprobado/suspense) es más popular que la predicción de la nota o puntuación final (Khan & Ghosh, 2021; Peña-Ayala, 2014). Diferentes algoritmos de clasificación han arrojado resultados notables en varios temas, es decir, bosque aleatorio (Gray & Perkins, 2019; Kumar & Singh, 2017), lógica difusa (Yildiz et al., 2013), agrupación de k-means (Sisovic et al., 2016), naive bayes (Kotsiantis et al., 2004), árbol de decisiones (Bunkar et al., 2012; Guruler et al., 2010), máquinas de vectores de soporte (Tekin, 2014), redes neuronales artificiales (Aydoğdu, 2019) y k-vecino más cercano (Nouri et al., 2019). Las diferencias en sus logros son bastante normales ya que el conjunto de datos de los estudiantes es diferente. Los mismos algoritmos pueden mostrar diferentes rendimientos para diferentes conjuntos de datos (Injadat et al., 2020; Kotsiantis et al., 2006). Además, cada algoritmo tiene algunos sesgos según el tipo de datos a los que se aplica, lo que puede dificultar la determinación del algoritmo universalmente aceptable. Por lo tanto, se recomienda utilizar modelos de aprendizaje en conjunto que combinen las predicciones de diferentes algoritmos para superar la capacidad de generalización, la robustez de un solo algoritmo de aprendizaje y hacer predicciones más precisas. (Kotsiantis et al., 2010). Los estudios deben llevarse a cabo con algoritmos de ML modernos donde los registros de aprendizaje sincrónicos y asincrónicos obtenidos de grandes grupos de muestra se incluyen en el proceso analítico (Corsatea & Walker, 2015; Korkmaz & Correia, 2019; Romero et al., 2013).

Este estudio tuvo como objetivo principal crear y optimizar un modelo de conjunto para predecir estudiantes en riesgo de fracaso académico durante la pandemia. Se aplicaron alrededor de cien ensayos en diferentes modelos de conjuntos que combinan análisis cuadrático discriminante (QDA), árbol de decisión (DT), bosque aleatorio (RF), árboles adicionales (ET), regresión logística (LR) y red neuronal artificial (ANN) algoritmos de clasificación. En este estudio se buscaron respuestas a las siguientes preguntas:

- ¿Qué modelo de conjunto es el mejor para predecir los estudiantes en riesgo de fracaso académico durante la pandemia?
- ¿Qué característica(s) de los estudiantes afectan el desempeño predictivo?

Este manuscrito está organizado de la siguiente manera: antecedentes en la sección "Antecedentes", metodología de la investigación en la sección "Metodología", resultados en la sección "Resultados", discusión en la sección "Discusión" y conclusiones y sugerencias en "Conclusión y sugerencia". sección.

Fondo

Predecir el desempeño de los estudiantes

DM es un proceso de descubrimiento de información que descubre las estructuras ocultas en grandes conjuntos de datos y obtiene información significativa para los tomadores de decisiones (Romero et al., 2013). ML se enfoca en el diseño y desarrollo de algoritmos que permiten a las computadoras desarrollar comportamientos y generar reglas basadas en datos empíricos (Singh & Lal, 2013). Reconoce automáticamente patrones complejos basados en datos pasados o actuales. Predice cuál será el valor de una característica objetivo en un gran volumen de datos (Singh & Lal, 2013). En los últimos años, los algoritmos de DM y ML se están utilizando ampliamente en educación, finanzas, marketing, atención médica, ingeniería y seguridad para aumentar su eficiencia y calidad. Estos algoritmos se pueden usar de manera eficiente en la educación superior para el descubrimiento de patrones de estudiantes, la automatización, el modelado de estudiantes y la predicción del rendimiento académico (Adejo y Connolly, 2017).

Durante la pandemia de COVID-19, el uso de sistemas de gestión de aprendizaje (LMS) en los sistemas de educación a distancia se ha incrementado exponencialmente, lo que ha producido grandes datos educativos. Sin embargo, el análisis manual de estos datos no es posible (Romero et al., 2008). Con la aplicación de algoritmos DM o ML para facilitar el análisis de datos educativos, han surgido dos nuevos campos de estudio, a saber, Minería de datos educativos (EDM) y Análisis de aprendizaje (LA). EDM & LA intersecta informática, educación y estadística (Romero & Ventura, 2013, 2020). Los temas principales de LA son la predicción del rendimiento, el apoyo a la decisión de profesores y alumnos, la detección de patrones de comportamiento y el modelado del alumno, y la predicción de la deserción escolar (Du et al., 2021). Los beneficios de LA para la educación incluyen una mayor participación de los estudiantes, mejores resultados de aprendizaje, identificación de estudiantes en riesgo, brindar retroalimentación en tiempo real y personalización del aprendizaje (Banihashem et al., 2018). EDM se enfoca en desarrollar modelos para mejorar la experiencia de aprendizaje y la efectividad institucional (Dutt et al., 2017; Hussain et al., 2021). En el estrechamente relacionado EDM & LA (Siemens & Baker, 2012) tiene como objetivo comprender y optimizar el proceso de aprendizaje (Gašević et al., 2016). Por tanto, la predicción del rendimiento de los alumnos tiene un lugar importante en los estudios realizados en estos campos (Banihashem et al., 2018; Du et al., 2021; Peña-Ayala, 2014; Romero & Ventura, 2020). Más específicamente, la predicción de estudiantes en riesgo de reprobar un curso (problema de clasificación) y la predicción de las calificaciones finales de los estudiantes (problema de regresión) son dos áreas de estudio comúnmente estudiadas.

Estudios previos relacionados con EDM y LA se han aplicado con éxito para predecir el rendimiento académico de los estudiantes (Aydoğdu, 2019; Bunkar et al., 2012; Gray & Perkins, 2019; Kotiantis et al., 2004; Kumar & Singh, 2017; Nouri et al., 2019; Sisovic et al., 2016; Tekin, 2014; Yildiz et al., 2013). Según Peña-Ayala (2014), el 60% de los artículos de investigación de EDM han utilizado el enfoque predictivo de DM. Del mismo modo, Shahiri et al. (2015) revisaron la predicción del desempeño de los estudiantes utilizando técnicas de DM y encontraron que el promedio de calificaciones (GPA) acumulativo y las evaluaciones internas son los atributos más frecuentes. También descubrieron que los árboles de decisión (DT) y las redes neuronales artificiales (ANN) eran las técnicas de DM más utilizadas para predecir el rendimiento de los estudiantes. Del mismo modo, Abu Saa et al. (2019) revisaron y analizaron 36 artículos de investigación de 2009 a 2018 y encontraron DT, naïve Bayes (NB) y ANN como los algoritmos de DM más comunes para predecir y clasificar los factores de los estudiantes. Los factores que afectan el desempeño del estudiante se encontraron como las calificaciones previas de los estudiantes, el desempeño en clase, la actividad de aprendizaje electrónico, la demografía de los estudiantes y la información social. Tomasevic et al. (2020) realizó un análisis exhaustivo

para comparar las técnicas de aprendizaje automático supervisado y encontró que las ANN son las mejores al alimentar los datos de participación de los estudiantes y los datos de rendimiento anteriores para las tareas de clasificación y regresión. No encontraron ninguna influencia de la demografía en las predicciones.

Estudios recientes han intentado identificar el mejor algoritmo de clasificación para predecir el rendimiento de los estudiantes (Akçapınar et al., 2019; Kotsiantis et al., 2004; Nouri et al., 2019). Por otro lado, Helal et al. (2018) se centró en diferentes algoritmos de clasificación para predecir el rendimiento de los estudiantes teniendo en cuenta la heterogeneidad de los estudiantes. Un discurso general es que todos trabajan juntos; ningún método individual muestra un rendimiento superior, pero los algoritmos basados en reglas, como los DT, proporcionaron la mayor capacidad de interpretación. Iatrellis et al. (2021) trató de adoptar un enfoque de aprendizaje automático de dos fases mediante la explotación de técnicas de aprendizaje tanto supervisadas como no supervisadas para predecir los resultados de los estudiantes donde los modelos consolidados produjeron predicciones relativamente precisas.

Eventualmente, predecir el desempeño de los estudiantes se ha convertido en una tarea desafiante. Los métodos de predicción existentes aún son insuficientes para predecir el desempeño de los estudiantes en las instituciones de educación superior (Abu Saa et al., 2019). Por lo tanto, había un claro requerimiento de métodos más avanzados que predijeran a los estudiantes en riesgo y determinaran qué características afectan los resultados de los estudiantes; que nos ha llevado a realizar este estudio.

aprendizaje conjunto

Se recomienda el uso de aprendizaje por conjuntos para aumentar la estabilidad de un solo algoritmo de aprendizaje y su precisión de predicción (Dietterich, 2000). El modelo de aprendizaje en conjunto se basa en un meta-algoritmo que combina el mismo o diferentes tipos de modelos entrenados individualmente para generar una predicción final (Kapucu & Cubukcu, 2021).

El aprendizaje de conjunto se puede clasificar según la variedad de algoritmos básicos de aprendizaje incluidos en el modelo creado y la forma en que se crea el modelo. En primer lugar, el aprendizaje conjunto puede ser heterogéneo u homogéneo según la variedad de algoritmos de aprendizaje incluidos en el modelo. En el modelo de conjunto heterogéneo, los mismos datos de entrenamiento se aplican a diferentes algoritmos de aprendizaje o a los mismos algoritmos con diferentes configuraciones de parámetros. En el homogéneo, los datos de entrenamiento originales se dividen en diferentes subconjuntos de datos y se aplican al mismo algoritmo de aprendizaje por el número de subconjuntos de datos (Wang et al., 2018). En segundo lugar, el aprendizaje conjunto se puede dividir en métodos de promediación y métodos de refuerzo. En los métodos de promediación, los modelos de aprendizaje se crean de forma independiente (Kapucu & Cubukcu, 2021). Las predicciones producidas por estos modelos básicos se promedian para reducir la varianza. En los métodos de impulso, los datos de entrenamiento originales se dividen en subconjuntos aleatorios. Las muestras del mismo algoritmo de aprendizaje básico seleccionado se entrenan con estos subconjuntos. Las predicciones obtenidas se combinan y se genera una predicción final. Los métodos de impulso se pueden aplicar con cualquier algoritmo de aprendizaje. Sin embargo, normalmente funcionan mejor con algoritmos potentes y complejos.

Recientemente, los investigadores han utilizado varios modelos de conjunto para predecir el éxito de los estudiantes. Kotsiantis et al. (2010) tuvo como objetivo llenar el vacío entre la predicción empírica del rendimiento de los estudiantes y las técnicas existentes de ML en un entorno de educación a distancia.

Tey propuso un conjunto de clasificadores en línea que combina una versión incremental de los algoritmos NB, 1-NN y WINNOW utilizando la metodología de votación. Encontraron el algoritmo propuesto como el más apropiado para construir una herramienta de soporte de software.

Injadat et al. (2020) propusieron un enfoque sistemático basado en el índice de Gini y el valor p

para seleccionar un conjunto de alumnos adecuado a partir de una combinación de varios algoritmos de aprendizaje automático. Analizaron dos conjuntos de datos diferentes en dos etapas separadas de la impartición del curso (20% y 50% respectivamente). Los resultados experimentales mostraron que los modelos de conjunto logran una alta precisión con una baja tasa de falsos positivos (FPR) en todas las etapas para ambos conjuntos de datos.

Metodología

Este estudio utilizó un modelo de conjunto que combina varios algoritmos de clasificación supervisados utilizando la metodología de votación para predecir los estudiantes en riesgo de fracaso académico durante la pandemia. La capacidad predictiva de los estudiantes en el éxito del curso se midió considerando las características de la actividad sincrónica y asincrónica (conjunto de datos). El sistema de predicción de estudiantes en riesgo consta de los siguientes pasos: agregación y preparación de datos, preprocesamiento de datos, optimización de algoritmos de subaprendizaje y creación y optimización de modelos de aprendizaje en conjunto (Fig. 1).

Todas las fases se realizaron en la distribución de Anaconda (una plataforma libre independiente del sistema operativo) con Python 3.x. Las bibliotecas utilizadas en Python fueron Scikit-learn (algoritmos ML), Pandas (para importar, crear y manipular marcos de datos), NumPy (computación de matriz), Matplotlib y Seaborn (visualización de datos). Scikit-learn es una gran biblioteca basada en NumPy, SciPy y Matplotlib que involucra herramientas simples y eficientes para el análisis predictivo de datos.

El núcleo de NumPy es un código C bien optimizado. La flexibilidad de Python viene con la velocidad del código compilado.

Agregación y preparación de datos

Este estudio se realizó en una universidad estatal turca. El conjunto de datos consiste en los datos de un curso obligatorio de 15 semanas titulado "Tecnologías de la información" de todos los estudiantes de primer grado y obtenido al final del semestre de otoño de 2020. Este estudio incorpora datos de actividad obtenidos del LMS universitario (Moodle) y del software de gestión de conferencias (Adobe Connect) durante un semestre de curso. Los registros de Moodle y Adobe Connect tienen diferentes actividades y recursos para los estudiantes. La Tabla 1 representa los significados de los diferentes atributos de actividad. Para realizar experimentos supervisados con estos conjuntos de datos, se eligió el resultado de la puntuación GPA (aprobado/reprobado) del curso como columna objetivo como preferencia general para determinar el rendimiento de los estudiantes. El puntaje GPA se obtiene del 40% del examen parcial y del 60% del examen final. El GPA se clasificó en dos categorías: (1) Aprobado: ≥ 50 , (2) Reprobado: < 50 .

El número de expedientes de alumnos fue de 2.116. Las características del conjunto de datos consisten tanto en variables góricas (grado, género) y variables numéricas como los números de descarga,

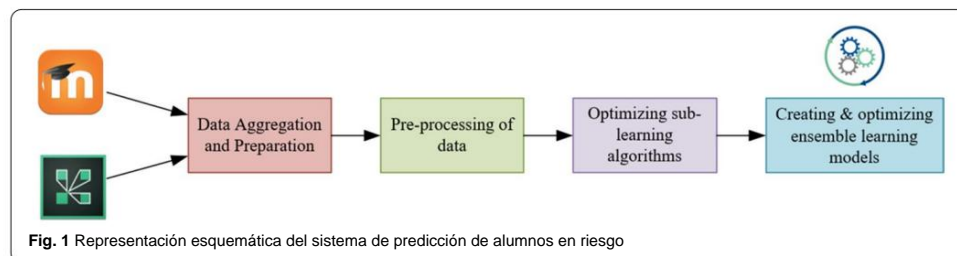


Tabla 1 Descripciones de funciones y rangos de datos del conjunto de datos

Nombre de la función	Descripción	Escribe	Valor rango
Género	Género (M/M)	Categorico	0–1
La licenciatura	Universidad (2 años) o facultad (4 años)	Categorico	0–1
Notas de lectura	Número de notas de clase descargadas en formato PDF	Numérico	0–203
Materiales	Número de otros materiales del curso descargados	Numérico	0–176
Video	Tiempo total dedicado a ver videos de cursos grabados (minutos)	Numérico	1–5988
Vivir	Tiempo total dedicado a las sesiones del curso en vivo (minutos)	Numérico	1-1295
Asistencia_en_vivo	Número total de asistencias en sesiones de cursos en vivo	Numérico	1–46
Examen	Puntuación de la prueba	Numérico	0–100

Tabla 2 Nuevos valores de asimetría para las características

Nombre de la función	Oblicuidad
Notas de lectura	0.3191
Materiales	0.2832
Video	ÿ 0,0094
Asistencia_en_vivo	0.1109
Vivir	0.0015

número de asistencia y tiempo total empleado en minutos. La puntuación de la prueba varía inherentemente entre 0 y 100, como se indica en la Tabla 1.

Preprocesamiento de datos

Los datos de los estudiantes se extrajeron de Moodle y Adobe Connect, se combinaron y filtraron de acuerdo con ciertos criterios. Se borraron los registros que contenían valores inconsistentes. Por ejemplo, las muestras que tienen una gran cantidad de sesiones de cursos en vivo pero un tiempo de asistencia muy bajo se eliminaron del conjunto de datos. Como resultado, en este estudio se evaluaron los registros del conjunto de datos de 2.045 estudiantes.

En la primera columna de la Tabla 2, las características (Lecture_notes, Materials, Video, Live_asistencia y en vivo) que no muestran la distribución normal en el conjunto de datos. Estas características se normalizaron utilizando el método de linealización porcentual. Como resultado, las características relevantes (Tabla 2) y el valor de asimetría se acercaron a cero y obtuvieron una forma de distribución casi normal (Fig. 2).

Lo último antes de crear el modelo de conjunto para determinar el desempeño de los estudiantes es la selección de funciones. En este paso, se determinaron y seleccionaron las características más adecuadas en el conjunto de datos en el entrenamiento de todos los algoritmos de aprendizaje candidatos para crear el modelo de conjunto final. Las funciones en el módulo sklearn.feature_selection se usaron para seleccionar características en conjuntos de muestra para mejorar las puntuaciones de precisión de los predictores o para aumentar su rendimiento. Aquí se utilizó SelectKBest, que es un método de selección de características univariante. SelectKBest elimina todas las funciones excepto las k de mayor puntuación. La selección de características univariante funciona seleccionando las mejores características de acuerdo con pruebas estadísticas univariadas. Puede verse como un paso de preprocesamiento antes de entrenar algoritmos. La Tabla 3 muestra las puntuaciones de las pruebas estadísticas univariadas de SelectKBest de las características en orden descendente.

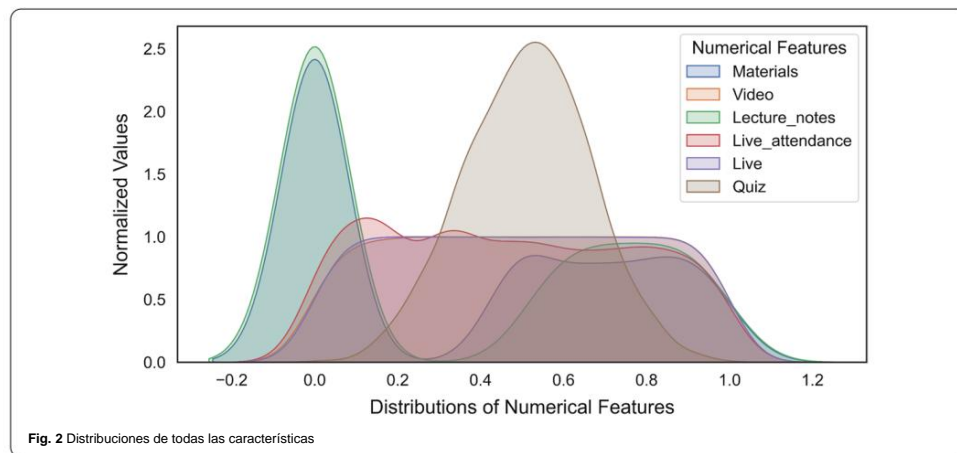


Tabla 3 Puntuaciones de la prueba SelectKBest de las funciones en el conjunto de datos de los estudiantes

Nombre de la función	Puntaje
Examen	24.49
La licenciatura	20.96
Notas de lectura	13.92
Materiales	9.30
Video	1.60
Género	1.53
Asistencia_en_vivo	0.48
Vivir	0.15

Está claro en la Tabla 3 que algunas funciones (como el cuestionario) son efectivas para predecir el desempeño de los estudiantes. Esta es una observación esperada; no hubo ningún cambio inesperado en este sentido durante la pandemia. Sin embargo, los efectos del video, la asistencia en vivo y las funciones en vivo, que son características de la educación sincrónica, en el desempeño de los estudiantes son sorprendentemente bajos. El género, por otro lado, no fue visto como un factor relevante en el éxito del curso. Se seleccionaron características distintas de aquellas cuyos factores de impacto son muy bajos de acuerdo con el valor de puntuación de la prueba estadística (Gender, Live_attendance y Live) para entrenar el modelo.

Optimización de algoritmos de subaprendizaje

Este estudio utilizó un modelo de aprendizaje de conjunto que incluye aumento de gradiente (GB), análisis discriminante cuadrático (QDA), árbol de decisión (DT), bosque aleatorio (RF), árboles adicionales (ET), regresión logística (LR) y red neuronal artificial (ANN) clasificadores como algoritmos candidatos. En la selección de conjuntos de algoritmos, la clasificación generalmente dio resultados exitosos donde cada uno miró el problema de clasificación desde una perspectiva diferente.

Las características generales de los algoritmos seleccionados se discuten en las siguientes líneas.

GB se basa en la intuición de que el siguiente modelo mejor posible, cuando se combina con modelos anteriores, minimiza el error de predicción general. La idea básica es establecer resultados objetivo para este próximo modelo para minimizar el error. GB se puede utilizar tanto para la clasificación como para

regresión. QDA (McLachlan, 1992) es un algoritmo de clasificación basado en el método de discriminación estadística que utiliza una superficie de decisión cuadrática para separar muestras de dos o más clases de objetos o eventos. Los modelos RF (Breiman, 2001) y ET (Geurts et al., 2006) son formas especializadas del método de embolsado en las que utilizan solo DT como algoritmo de aprendizaje base. En los métodos de embolsado, los datos de entrenamiento originales se dividen en subconjuntos aleatorios. Se entrenan copias idénticas de los algoritmos básicos de aprendizaje seleccionados con estos subconjuntos. Se utiliza una combinación de estas predicciones para obtener una predicción final. Estos métodos se utilizan para reducir la varianza del algoritmo subyacente. Así también se reduce el sobreajuste, que es un problema de clasificación. En los modelos RF y ET, hay varios DT. Dan la predicción promediada de estos árboles combinados. A diferencia del RF, el modelo ET utiliza un conjunto de datos de entrenamiento completo en lugar de réplicas remuestreadas para hacer crecer los árboles y dividir los nodos del árbol al azar (Kapucu & Cubukcu, 2021). LR es un modelo estadístico que emplea una función logística para modelar una variable dependiente binaria. Un modelo logístico binario tiene una variable dependiente matemáticamente con dos valores potenciales, es decir, pasa/falla, indicado por 0 y 1. ANN es una estructura de procesamiento de información paralela y distribuida inspirada en el cerebro humano, que consta de unidades de procesamiento interconectadas (Russell & Norving, 2020). Esta estructura se compone de una combinación de células nerviosas artificiales (neuronas) creadas imitando las funciones de las células nerviosas biológicas.

Las neuronas artificiales son unidades y sistemas de información completamente inspirados en la naturaleza que pueden imitar la capacidad de aprendizaje del cerebro humano. Una red neuronal es una estructura de red creada por la unión de neuronas entre sí (Haykin, 2008).

Al determinar si un algoritmo de aprendizaje es adecuado para el campo y los datos utilizados, el modelo obtenido con el entrenamiento del algoritmo se prueba con un conjunto de datos que contiene ejemplos que no se han encontrado anteriormente. Durante la prueba de un modelo, la fuente de datos sigue siendo la misma, mientras que las muestras deben ser nuevas. Por esta razón, es el método preferido para dividir el conjunto de datos en propiedad en partes de entrenamiento y prueba. En casos de escasez de muestras, el éxito de generalización en los datos puede disminuir debido al entrenamiento incompleto del algoritmo de aprendizaje. Por lo tanto, el método de división de entrenamiento/prueba es más adecuado para grandes conjuntos de datos. Para superar este problema, se propuso el método de validación cruzada (CV) (Allen, 1974; Geisser, 1975). El CV es un método de división de prueba de entrenamiento que divide el conjunto de datos original en (K) piezas. Cada pieza se utiliza tanto en las partes de entrenamiento como de prueba del modelo creado.

La función de búsqueda en cuadrícula se utilizó para optimizar los algoritmos candidatos. Esta función realiza una búsqueda iterativa para encontrar los valores de hiperparámetro óptimos para un algoritmo de aprendizaje en particular. Usando CV durante el proceso de búsqueda, la función informa los mejores parámetros candidatos y las precisiones de predicción obtenidas con esta configuración como resultado. Mientras se optimizaba el modelo dentro del alcance de la investigación, se llevó a cabo un CV diez veces mayor para los hiperparámetros (Tabla 4) y la búsqueda en cuadrícula observando el rendimiento de la especificidad.

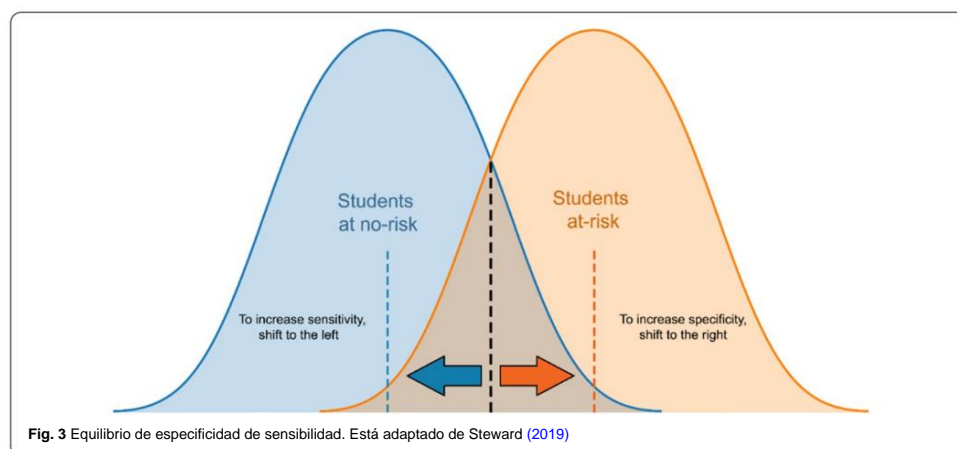
La métrica de rendimiento de especificidad puede hacer una clasificación en la que se puede predecir con mayor precisión a los estudiantes con riesgos académicos, que es el objetivo principal, sin ignorar a los estudiantes exitosos (Fig. 3). Al desplazarse hacia la derecha, la proporción de FP disminuye, lo que significa una especificidad mejorada.

Esta medida de rendimiento establecida se denomina comúnmente tasa negativa verdadera (TNR) y selectividad. Un modelo probado de esta manera logra un resultado de TN alto con un FP bajo para obtener

Tabla 4 Parámetros de búsqueda en cuadrícula para cada algoritmo candidato

Algoritmos Espacio de hiperparámetros para mejorar el rendimiento de los algoritmos	
DT	Medida de impureza=['gini', 'entropía'] & Estrategia dividida=['mejor', 'aleatorio'] & Profundidad máxima=[Ninguno, 3] & Funciones máximas = ['auto', 'sqrt', 'log2'] & Peso de clase = ['equilibrado', Ninguno]
RF	Medida de impureza=['gini', 'entropía'] & Bootstrap=[False, True] & Max depth=[Ninguno, 3] & Warm start=[False, True] & Class weight=['balanced', 'balanced_subsample'] & Número de árboles=[100, 200, 300, 400]
Y	Medida de impureza=['gini', 'entropía'] & Profundidad máxima=[Ninguno, 3] & Inicio cálido=[Falso, Verdadero] & Peso de clase=['equilibrado', 'submuestra_equilibrada'] & Número de árboles=[100, 200, 300, 400]
LR	Algoritmo de optimización=['newton-cg', 'lbfgs', 'liblinear', 'sag', 'saga'] & Inverso de fuerza de regularización=[0.1, 1, 10, 100] & Peso de clase=['equilibrado', Ninguno] & Dual=[Falso, Verdadero] & Fit intercept=[False, True] & Tol=[0.001, 0.01] & Warm start=[False, True]
ES	Medida de impureza=['friedman_mse', 'mse', 'mae'] & Profundidad máxima=[Ninguno, 3] & Número de árboles=[100, 200, 300]
ANA	Alfa=[0.1, 0.01, 0.001] & Activación=['relu', 'logística'] & Detención anticipada=[Verdadero, Falso] & Número de capas ocultas=[200, 300, 400]
QDA	*No hay configuraciones de parámetros disponibles en la implementación de Scikit-learn

Árbol de decisión (DT), Bosque aleatorio (RF), Árboles adicionales (ET), Regresión logística (LR), Aumento de gradiente (GB), Neural artificial Red (ANN) y Análisis Discriminante Cuadrático (QDA)



una puntuación alta La versión Scikit-learn del algoritmo QDA se usó con su configuración predeterminada, ya que no hay hiperparámetros que se puedan ajustar para el rendimiento. Las métricas preferidas y utilizadas con frecuencia en la literatura al medir el rendimiento de clasificación del modelo se presentan brevemente a continuación. Estas métricas utilizan valores verdaderos positivos (TP), verdaderos negativos (TN), falsos positivos (FP) y falsos negativos (FN) derivados de la matriz de confusión (Tabla 5).

Creación y optimización de modelos de aprendizaje de conjunto

Para encontrar el modelo de conjunto más exitoso, primero, todos los algoritmos de aprendizaje se combinan en conjuntos dobles, triples y cuádruples. De esta forma, los modelos creados intentan predecir mejor la clase objetivo combinando conceptualmente diferentes modelos de sub-aprendizaje. Dentro del alcance de este estudio, se crearon modelos de conjuntos candidatos en los que cada modelo de aprendizaje en el grupo tiene el mismo peso y los mismos derechos de voto (clasificador de votación). Por lo tanto, todos los modelos de conjunto creados en este paso se probaron utilizando un CV de diez veces con una métrica de rendimiento de precisión equilibrada. En el ranking obtenido, los más exitosos

Tabla 5 Métricas de rendimiento de clasificación

Métrico	Explicación	Fórmula
Exactitud	¿Con qué frecuencia son correctas las predicciones del modelo?	$\frac{TP+TN}{TP+TN+FN+FP}$
Precisión	Cuando el modelo predice positivo, ¿con qué frecuencia es correcto? $TP / (TP+FP)$	
Sensibilidad (recuperación) o TPR	Cuando en realidad es positivo, ¿con qué frecuencia predice correctamente? $TP / (TP+FN)$	
Especificidad o TNR	Cuando en realidad es negativo, ¿con qué frecuencia predice correctamente? $TN / (TN+FP)$	
Precisión equilibrada	La media de sensibilidad y especificidad.	$(TPR+TNR) / 2$
Medida F	La media armónica de precisión y sensibilidad.	$2 \times \text{Precisión} \times \text{Recuperación} / (\text{Precisión} + \text{Recuperación})$

Se determinaron los modelos de conjuntos candidatos. Esta métrica asegura que, entre todos los modelos de conjunto candidatos, en especial aquellos que pueden predecir ambas clases objetivo con alta precisión, estén en la parte superior.

Finalmente, después de encontrar los modelos de conjuntos candidatos más exitosos, se optimizaron. Para ello, y para determinar el mejor de ellos como modelo de conjunto final, se realizó una búsqueda iterativa en los diferentes tipos de voto y ponderaciones de derechos de voto enumerados en la Tabla 6. En este proceso de búsqueda, los modelos de conjunto candidatos más exitosos se probaron con estas diferentes configuraciones utilizando una métrica de rendimiento de especificidad con un CV diez veces mayor.

Resultados

Optimización de algoritmos de subaprendizaje

Las estadísticas de prueba de los algoritmos de aprendizaje utilizados en el estudio se dan en la Tabla 7. En consecuencia, los modelos más exitosos en los resultados de TN, que muestran las predicciones correctas de estudiantes que no tuvieron éxito (reprobaron) en el conjunto de datos de prueba, fueron RF, LR y DT, mientras que

Tabla 6 Configuraciones de optimización de los modelos de conjuntos candidatos más exitosos

Parámetros	Opciones
tipo de votacion	'suave duro'
Pesos	Ninguno, [1,1,2], [1,2,1], [2,1,1], [1,2,2], [2,2,1], [2,1,2]

Tabla 7 Estadísticas de prueba de los algoritmos individuales optimizados con grid-search

Modelos	Tamano	FP	FN	TP
Aumento de gradiente (GB)	91	54	38	431
Análisis Discriminante Cuadrático (QDA)	81	64	25	444
Árboles adicionales (ET)	112	33	107	362
Árbol de decisión (DT)	120	25	250	219
Bosque aleatorio (RF)	125	20	100	369
Regresión Logística (LR)	121	24	89	380
Red neuronal artificial (ANN)	86	59	26	443

Verdaderos positivos (TP), Verdaderos negativos (TN), Falsos positivos (FP) y Falsos negativos (FN)

Tabla 8 Puntuaciones de rendimiento de los algoritmos individuales optimizados con grid-search

	ES	QDA	Y	DT	RF	LR	ANA
Especificidad o TNR	0.6276	0.5586	0.7724	0.8276	0.8621	0.8345	0.5931
Sensibilidad o TPR	0.9190	0.9467	0.7719	0.4670	0.7868	0.8102	0.9446
Precisión	0.8887	0.8740	0.9165	0.8975	0.9486	0.9406	0.8825

Los caracteres en cursiva muestran las puntuaciones más exitosas de cada fila

Aumento de gradiente (GB), análisis cuadrático discriminante (QDA), árboles adicionales (ET), árbol de decisión (DT), bosque aleatorio (RF), Regresión logística (LR) y red neuronal artificial (ANN)

Tabla 9 Estadísticas de prueba de los cinco mejores modelos de conjuntos candidatos con optimización

Conjunto candidato y su submodelos	Tamano	FP	FN	TP
ET+RF+LR	131	14	110	359
QDA + LR	121	24	89	380
GB+ET+RF+LR	121	24	91	378
GB+DT+LR	117	28	90	379
GB + DT + LR + ANN	118	27	91	378

Los caracteres en cursiva muestran la predicción negativa verdadera más alta.

Aumento de gradiente (GB), análisis cuadrático discriminante (QDA), árboles adicionales (ET), árbol de decisión (DT), bosque aleatorio (RF), Regresión logística (LR) y red neuronal artificial (ANN)

los modelos con menos éxito fueron QDA, ANN y GB. El modelo ET tuvo un éxito moderado.

La Tabla 8 muestra las puntuaciones de rendimiento calculadas en el rango de [0–1] en función de los resultados de predicción obtenidos por los modelos en el conjunto de datos de prueba. Los modelos más exitosos en los valores de especificidad que muestran las tasas de estimación correctas de los estudiantes reprobados en el conjunto de datos fueron RF, LR y DT, mientras que los modelos con menos éxito fueron QDA, ANN y GB. El modelo ET tuvo un éxito moderado. En los valores de sensibilidad, que muestran las tasas de estimación correctas de los alumnos que aprobaron el curso en el conjunto de datos, los modelos más exitosos fueron QDA, ANN y GB, mientras que otros modelos mostró menor éxito.

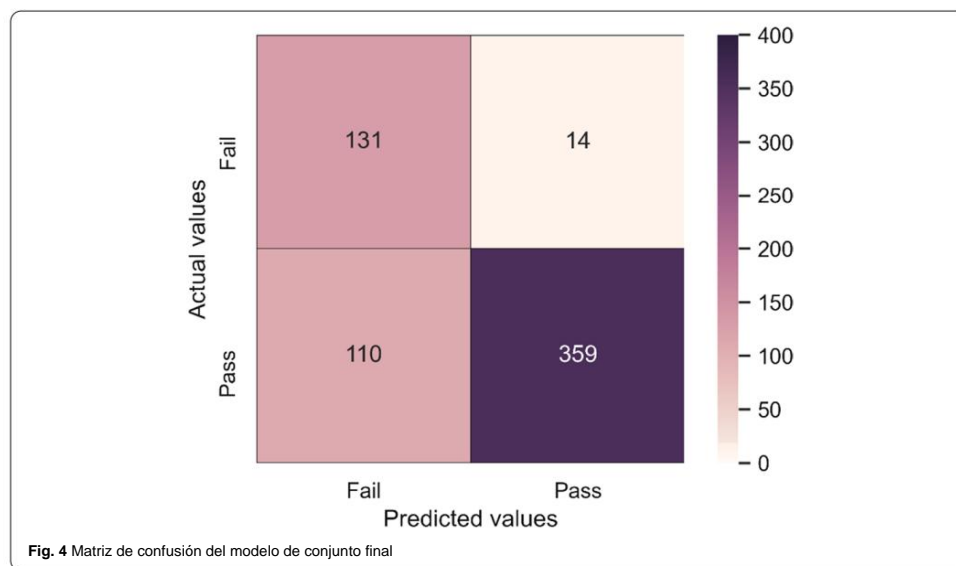
Creación y optimización del modelo de aprendizaje de conjunto

Se crearon y probaron un total de 91 modelos de conjuntos candidatos utilizando combinaciones dobles, triples y cuádruples de siete algoritmos de clasificación. Entre estos modelos, las estadísticas de prueba de los 5 modelos de conjunto más exitosos que pueden predecir ambas clases de objetivos (aprobado o reprobado) junto con alta precisión mediante el uso de métricas de rendimiento de precisión equilibrada se dan en la Tabla 9.

Al final del proceso de búsqueda, el primer modelo candidato formado por los submodelos ET, RF, LR fue seleccionado como el modelo de conjunto final. En el modelo de conjunto final, el tipo de votación "dura" y los pesos del modelo [1,2,1] se determinaron como los mejores parámetros para la optimización.

La figura 4 muestra la matriz de error del modelo de conjunto final seleccionado.

Las primeras tres filas de la Tabla 10 muestran las puntuaciones de rendimiento de todos los modelos en el conjunto de datos de prueba. En la última fila de la tabla, el promedio de las precisiones de estimación realizadas

**Tabla 10** Puntuaciones de rendimiento de algoritmos individuales y modelo de conjunto

	ES	QDA	Y	DT	RF	LR	ANA	OLMO
especificidad	0.6276	0.5586	0.7724	0.8276	0.8621	0.8345	0.5931	0.9034
Sensibilidad	0.9190	0.9467	0.7719	0.4670	0.7868	0.8102	0.9446	0.7655
Precisión	0.8887	0.8740	0.9165	0.8975	0.9486	0.9406	0.8825	0.9625
CV—Especificidad	0.6099	0.5768	0.7841	0.8445	0.8425	0.8362	0.5768	0.8861

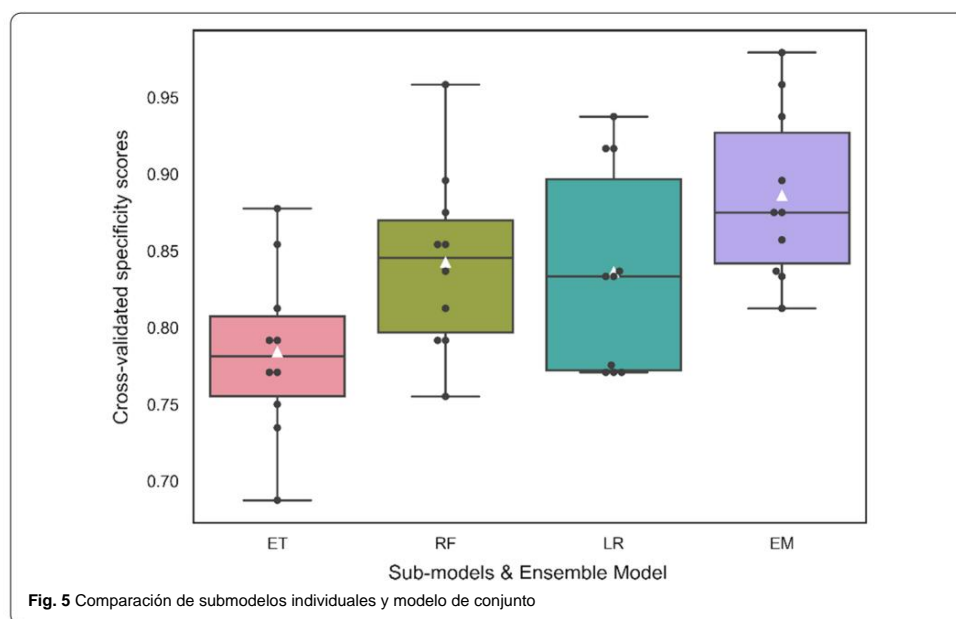
Aumento de gradiente (GB), análisis cuadrático discriminante (QDA), árboles adicionales (ET), árbol de decisión (DT), bosque aleatorio (RF), Regresión logística (LR), red neuronal artificial (ANN), modelo de aprendizaje de conjunto (ELM)

Se proporciona el conjunto de datos completo con un CV de diez veces, considerando la métrica de rendimiento de especificidad de todos los modelos. Los caracteres en cursiva muestran las puntuaciones más exitosas de cada fila. Dado que este estudio tuvo como objetivo predecir los estudiantes en riesgo, los valores de especificidad-TNR (primera fila) que se muestran en la Tabla 10 son más importantes. El modelo de conjunto final logró la puntuación más alta tanto en el conjunto de datos de prueba únicamente como en todo el conjunto de datos con CV de acuerdo con la métrica de rendimiento de especificidad. Al mismo tiempo, fue el modelo que logró la puntuación más alta en la métrica de rendimiento de precisión.

La figura 5 muestra las puntuaciones de especificidad validadas de forma cruzada de los algoritmos individuales y el modelo de conjunto. La comparación visual de diagrama de caja de las puntuaciones de especificidad para los submodelos ET, RF, LR y el modelo de conjunto optimizado se obtuvo con un CV de diez veces. Los puntos negros en cada diagrama de caja representan los puntajes obtenidos en cada pliegue del CV, mientras que los triángulos blancos representan el promedio de estos puntajes. Como puede verse, la puntuación media obtenida por el modelo de conjunto final en el extremo derecho del gráfico es más alta que la de los tres submodelos.

Discusión

La pregunta es, ¿qué modelo de conjunto es el mejor para predecir los estudiantes en riesgo de fracaso académico durante la pandemia? Para responder a esta pregunta, se crearon y probaron 91 modelos de conjuntos candidatos con combinaciones dobles, triples y cuádruples.



a partir de algoritmos de clasificación que utilizan el aumento de gradiente, el análisis discriminante cuadrático, el árbol de decisión, el bosque aleatorio, los árboles adicionales, la regresión logística y la red neuronal artificial. Los resultados mostraron que el modelo de conjunto, que consiste en combinaciones de algoritmos de clasificación de árboles extra (ET), bosque aleatorio (RF) y regresión logística (LR), es el mejor para predecir los estudiantes en riesgo después de la optimización. La especificidad de este modelo, es decir, la puntuación TNR es del 90,34 %. En estudios anteriores, los mejores algoritmos utilizados para predecir el rendimiento de los estudiantes son Random Forest (Gray & Perkins, 2019; Kumar & Singh, 2017), Fuzzy Logic (Yildiz et al., 2013), K-means Clustering (Sisovic et al., 2016), Naive Bayes (Kotsiantis et al., 2004), Decision Tree (Bunkar et al., 2012), Support Vector Machines (Tekin, 2014), Artificial Neural Net work (Aydođdu, 2019) y kNN (Nouri et al. ., 2019). Sin embargo, este estudio reveló el modelo de conjunto que da el mejor resultado de predicción con la combinación de diferentes algoritmos de clasificación en lugar de determinar el algoritmo individual que mejor predice el rendimiento de los estudiantes. En un estudio similar de Kotsiantis et al. (2010), se encontró que el modelo de conjunto que consiste en la combinación de los algoritmos Naive Bayes, 1-NN y WINNOWER predice mejor el rendimiento de los estudiantes. Los resultados muestran que los modelos de conjunto pueden tener más éxito en la predicción del desempeño de los estudiantes, especialmente en la predicción de estudiantes en riesgo. Por esta razón, se recomienda utilizar el enfoque de aprendizaje conjunto en los estudios que se realizarán para la predicción del desempeño de los estudiantes.

Otra pregunta es ¿qué características de los estudiantes afectan el desempeño predictivo?

Con respecto a esta pregunta, los resultados mostraron que los factores más importantes que afectan el rendimiento de la predicción son la puntuación del examen, el grado, la cantidad de notas de lectura descargadas, la cantidad de otros materiales del curso descargados y el tiempo total dedicado a ver los videos grabados del curso. La eficacia de las características de los estudiantes (como la calificación y el puntaje del examen) utilizadas en esta investigación para predecir el desempeño de los estudiantes es obvia. Esto es

una observación esperada. Aquí no se ha visto el efecto de la pandemia. Los estudios muestran que existe una relación entre el uso de Moodle y el desempeño de los estudiantes (Coursate & Walker, 2015). Sin embargo, los hallazgos sobre qué tipos de interacción afectan el desempeño de los estudiantes son contradictorios. Muestra que un modelo de regresión múltiple puede predecir el 52% de la varianza en los puntajes de los exámenes finales de los estudiantes al enviar y leer mensajes, contribuyendo a la producción de contenido, la cantidad de cuestionarios completados y archivos examinados (Zacharis, 2015). En otro estudio (Aydoğdu, 2019), se concluyó que el número de asistencias a las lecciones en vivo, el número de asistencias a las lecciones archivadas y el tiempo dedicado al contenido fueron más efectivos que otras variables para predecir el desempeño de los estudiantes. En otro estudio (You, 2016), se determinó que la entrega tardía de tareas, el número de sesiones asistidas o el inicio de sesión en el curso y la lectura del paquete de información del curso son variables importantes para predecir el éxito del curso. En consecuencia, el uso de Moodle tiene un efecto determinante en el rendimiento de los alumnos. Sin embargo, qué tipos de interacción son más efectivos para predecir el rendimiento aquí pueden variar según el contexto.

En esta investigación, sorprende que las características live (tiempo total dedicado a las sesiones del curso en vivo) y live_attendance (número total de asistencias a las sesiones del curso en vivo) obtenidas no son efectivas. Paralelamente a este hallazgo, las puntuaciones de los alumnos que han asistido a las sesiones en directo no son diferentes a las de los que han visto los videos grabados (Nieuwoudt, 2020). Por lo tanto, las sesiones en línea sincrónicas no son efectivas para predecir el rendimiento de los estudiantes (Gašević et al., 2016). Un informe anterior también mostró un éxito similar de los estudiantes de ambos grupos que habían aprendido de forma sincrónica y asincrónica (Olson & McCracken, 2015). Por el contrario, en otro estudio, se encontró que las sesiones de cursos en vivo fueron significativas para predecir el desempeño de los estudiantes (Aydoğdu, 2019; You, 2016). Las decisiones estratégicas tomadas por la administración universitaria pueden haber afectado las preferencias de los estudiantes en el presente estudio. No se requirió que los estudiantes asistieran a las sesiones en vivo, considerando la posibilidad de que algunos estudiantes no tengan acceso a una computadora o conexión a Internet. Les bastó con ver los videos grabados de las lecciones más tarde a sus anchas. Entonces, la mayoría de los estudiantes han visto las lecciones de forma asincrónica. Por otro lado, la mayoría de los estudiantes pueden haber perdido interés en las lecciones en vivo. En resumen, las sesiones del curso en vivo resultaron ineficaces; sin embargo, los videos grabados de las sesiones fueron efectivos para predecir el riesgo de los estudiantes. Con base en estos hallazgos, los administradores educativos y los formuladores de políticas deberían considerar el uso de software gratuito de administración de conferencias para usar de manera eficiente los recursos limitados. Para aumentar el éxito de los estudiantes, se recomienda que los educadores graben las sesiones en vivo y compartan estas grabaciones con los estudiantes.

En este estudio, se obtuvieron resultados con cinco características en el conjunto de datos de entrenamiento. Es necesario mejorar la capacidad de generalización del modelo establecido para una mayor precisión de predicción. Por esta razón, se recomienda agregar nuevas características efectivas para entrenar el modelo en estudios similares. Además, la cantidad de observaciones se puede aumentar para eliminar la dependencia de los comportamientos específicos del curso, del instructor y del estudiante en el conjunto de datos. Además, usar datos obtenidos de más de un semestre y volver a entrenar el modelo varias veces dentro de un período para aumentar la capacidad de generalización del modelo y, por lo tanto, hacer una predicción exitosa para el próximo semestre puede aumentar el rendimiento de la predicción. No obstante, un conjunto de características obtenidas del aprendizaje

El sistema de gestión o el software de gestión de conferencias puede tener un impacto significativo en el rendimiento de los estudiantes en un curso específico, mientras que no es tan beneficioso en otro (Gašević et al., 2016).

Como es sabido, el éxito de la educación a distancia depende en gran medida de la calidad de la comunicación e interacción con el estudiante. Por lo tanto, los administradores y los responsables de la formulación de políticas deben alentar a los instructores a usar entornos de aprendizaje sincrónicos y asincrónicos de manera efectiva. Los instructores también deben alentar a los estudiantes a usar de manera efectiva estos entornos educativos. Dado que, al aumentar la capacidad de generalización de los modelos entrenados con datos que contienen muchas características efectivas basadas en la comunicación y la interacción de abajo hacia arriba, los formuladores de políticas pueden tomar decisiones basadas en datos que pueden aumentar la calidad de la educación superior.

Para finalizar, este estudio desarrolla un modelo de conjunto y prueba datos obtenidos de más de dos mil estudiantes en lugar de grupos pequeños e identifica predictores importantes de estudiantes en riesgo de fracaso académico durante la pandemia.

Conclusión y sugerencias

El modelo de aprendizaje en conjunto proporciona una alta precisión en muchas áreas. Su uso como un nuevo enfoque se está extendiendo rápidamente en la educación. Este estudio concluyó que el enfoque del modelo de aprendizaje en conjunto, que consiste en combinaciones de algoritmos de clasificación de árboles extra (ET), bosque aleatorio (RF) y regresión logística (LR), es el mejor para predecir el riesgo de fracaso académico de los estudiantes. El enfoque propuesto en este estudio ayudará a los administradores, instructores y legisladores a desarrollar nuevas políticas e intervenciones educativas en la educación superior, ya que predice a los estudiantes en riesgo con gran precisión. Con estas políticas e intervenciones educativas por determinar, los estudiantes en riesgo pueden recibir el apoyo y la retroalimentación necesarios para evitar que se retrasen o fallen. Esta situación puede aumentar la calidad, la eficiencia y la eficacia de la educación en la educación superior.

Otro resultado del estudio es que la puntuación de la prueba, el grado, la cantidad de notas de lectura descargadas, la cantidad de descargas de otros materiales del curso y el tiempo total dedicado a ver los videos grabados del curso son efectivos para predecir los estudiantes en riesgo. Para predecir los estudiantes en riesgo con mayor precisión, se recomienda que los instructores y los diseñadores de instrucción presten atención a los cuestionarios, las características de los títulos de los estudiantes, las notas de clase compartidas y los videos grabados del curso mientras diseñan el contenido del curso. Por otro lado, el aprendizaje sincrónico durante la pandemia fue ineficaz para predecir los estudiantes en riesgo. Las razones de esta situación pueden ser investigadas en profundidad. No se incluyeron funciones como chat, wiki, foro, diccionario y cuestionarios en el sistema Moodle, la participación verbal de los estudiantes en el sistema Adobe Connect y los registros de mensajes de chat de los estudiantes con el instructor y otros estudiantes durante la lección. el estudio. En estudios posteriores, se sugiere probar los modelos de aprendizaje en conjunto creados con la combinación de diferentes algoritmos utilizando más datos sobre las características de los estudiantes y las actividades de aprendizaje sincrónicas y asincrónicas. Más específicamente, los modelos deben ser entrenados agregando nuevas funciones sin permitir que un modelo ajustado en exceso mejore la bondad del modelo.

Agradecimientos

No aplica.

Contribuciones de los autores

En nombre de todos los autores, el autor correspondiente declara que todos los autores han contribuido por igual. Todos los autores leyeron y aprobaron el manuscrito final.

Fondos

No aplica.

Disponibilidad de datos y materiales.

Los datos que respaldan los hallazgos de este estudio están disponibles a pedido.

Declaraciones**Conflicto de intereses**

Los autores declaran no tener conflictos de intereses.

Detalles del autor

¹ Departamento de Educación Informática y Tecnologías Educativas, Facultad de Educación, Universidad Muğla Sıtkı Koçman ciudad, Muğla, Turquía. ² Departamento de Informática, Universidad Muğla Sıtkı Koçman, Muğla, Turquía. ³ Departamento de Información Ingeniería de Sistemas, Facultad de Tecnología, Universidad Muğla Sıtkı Koçman, Muğla, Turquía.

Recibido: 24 mayo 2021 Aceptado: 22 septiembre 2021

Published online: 02 December 2021

Referencias

- Abu Saa, A., Al-Emran, M. y Shaalan, K. (2019). Factores que afectan el desempeño de los estudiantes en la educación superior: una revisión sistemática de las técnicas predictivas de minería de datos. *Tecnología, conocimiento y aprendizaje*, 24, 567–598. <https://doi.org/10.1007/s10758-019-09408-7>.
- Adejo, O. y Connolly, T. (2017). Un marco de sistema integrado para predecir el rendimiento académico de los estudiantes en instituciones de educación superior. *Revista internacional de informática y tecnología de la información*, 9(3), 149–157. <https://doi.org/10.5121/ijcsit.2017.93013>
- Akçapınar, G., Altun, A. y Aýkar, P. (2019). Uso de análisis de aprendizaje para desarrollar un sistema de alerta temprana para estudiantes en riesgo. *Revista Internacional de Tecnología Educativa en la Educación Superior*. <https://doi.org/10.1186/s41239-019-0172-z>
- Allen, DM (1974). La relación entre la selección de variables y la agregación de datos y un método para la predicción. *Tecnometría*, 16(1), 125–127. <https://doi.org/10.2307/1267500>
- Aydođdu, Ÿ. (2019). Predicción del rendimiento final de los estudiantes utilizando redes neuronales artificiales en entornos de aprendizaje en línea. *Educación y Tecnologías de la Información*, 25(3), 1913–1927. <https://doi.org/10.1007/s10639-019-10053-x>
- Banihashem, SK, Aliabadi, K., Pourroostaei Ardakani, S., Delaver, A. y Nili Ahmadabadi, M. (2018). Análisis de aprendizaje: una revisión sistemática de la literatura. *Revista Interdisciplinaria de Aprendizaje Virtual en Ciencias Médicas*. <https://doi.org/10.5812/ijvms.63024>
- Breiman, L. (2001). Bosques aleatorios. *Aprendizaje automático*, 45, 5–32. <https://doi.org/10.1023/A:1010933404324>
- Bunkar, K., Singh, Reino Unido, Pandya, B. y Bunkar, R. (2012). Minería de datos: Predicción para la mejora del desempeño de los estudiantes de posgrado utilizando la clasificación. *Conferencia Internacional IFIP sobre Redes de Comunicaciones Ópticas e Inalámbricas, WOCN*. <https://doi.org/10.1109/WOCN.2012.6335530>
- Corsate, BM y Walker, S. (2015). Oportunidades para datos de Moodle e inteligencia de aprendizaje en entornos virtuales. *Conferencia internacional IEEE de 2015 sobre sistemas inteligentes adaptativos y en evolución, EAIS 2015*, 1–6. doi: <https://doi.org/10.1109/EAIS.2015.7368776>
- Consejo de Educación Superior. (2020). Rueda de prensa. <https://www.yok.gov.tr/Sayfalar/Haberler/2020/universitelerde-uygul-pero-divulgacion-relacionada-con-la-educacion-remota.aspx>. Consultado el 21 de marzo de 2021.
- Dietterich, TG (2000). Métodos de conjunto en el aprendizaje automático. En: *Sistemas clasificadores múltiples*. MCS 2000. Apuntes de clase en Informática, vol 1857. https://doi.org/10.1007/3-540-45014-9_1
- Du, X., Yang, J., Shelton, BE, Hung, JL y Zhang, M. (2021). Una meta-revisión y análisis sistemáticos de la investigación de análisis de aprendizaje. *Comportamiento y tecnología de la información*, 40(1), 49–62. <https://doi.org/10.1080/0144929X.2019.1669712>
- Dutt, A., Ismail, MA y Herawan, T. (2017). Una revisión sistemática sobre la minería de datos educativos. *Acceso IEEE*, 5, 15991–16005. <https://doi.org/10.1109/ACCESO.2017.2654247>
- Gašević, D., Dawson, S., Rogers, T. y Gasevic, D. (2016). El análisis de aprendizaje no debe promover una talla única para todos: los efectos de las condiciones de instrucción en la predicción del éxito académico. *Internet y Educación Superior*, 28, 68–84. <https://doi.org/10.1016/j.iheduc.2015.10.002>
- Geisser, S. (1975). El método predictivo de reutilización de muestras con aplicaciones. *Revista de la Asociación Estadounidense de Estadística*, 70(350), 320–328. <https://doi.org/10.2307/2285815>
- Geurts, P., Ernst, D. y Wehenkel, L. (2006). Árboles extremadamente aleatorios. *Aprendizaje automático*, 63(1), 3–42. <https://doi.org/10.1007/s10994-006-6226-1>
- Gray, CC y Perkins, D. (2019). Utilizar el compromiso temprano y el aprendizaje automático para predecir los resultados de los estudiantes. *Computers & Education*, 131 (diciembre de 2018), 22–32. <https://doi.org/10.1016/j.compedu.2018.12.006>
- Guruler, H., Istanbulu, A. y Karahasan, M. (2010). Un nuevo sistema de análisis del rendimiento de los estudiantes utilizando el conocimiento descubrimiento en bases de datos de educación superior. *Informática y Educación*, 55(1), 247–254. <https://doi.org/10.1016/j.compe.2010.01.010>
- Haykin, S. (2008). *Redes neuronales y máquinas de aprendizaje* (3ª ed.). Educación Pearson.

- Helal, S., Li, J., Liu, L., Ebrahimi, E., Dawson, S., Murray, DJ y Long, Q. (2018). Predecir el rendimiento académico por teniendo en cuenta la heterogeneidad de los estudiantes. *Sistemas basados en el conocimiento*, 161, 134–146. <https://doi.org/10.1016/j.knosys.2018.07.042>
- Hussain, S., Gaftandzhieva, S., Maniruzzaman, M., Doneva, R. y Muhsin, ZF (2021). Análisis de regresión del estudiante aca Rendimiento demográfico usando aprendizaje profundo. *Educación y Tecnologías de la Información*, 26(1), 783–798. <https://doi.org/10.1007/s10639-020-10241-0>
- Iatrellis, O., Savvas, I., Fitsilis, P. y Gerogiannis, VC (2021). Un enfoque de aprendizaje automático de dos fases para predecir los resultados de los estudiantes. *Educación y tecnologías de la información*, 26(1), 69–88. <https://doi.org/10.1007/s10639-020-10260-x>
- Injadat, MN, Moubayed, A., Nassif, AB y Shami, A. (2020). Enfoque de selección de modelo de conjunto sistemático para educa minería de datos nacional. *Sistemas basados en el conocimiento*, 200, 105992. <https://doi.org/10.1016/j.knosys.2020.105992>
- Kapucu, C. y Cubukcu, M. (2021). Un método de aprendizaje de conjunto supervisado para el diagnóstico de fallas en cadenas fotovoltaicas. *Energía*, 227, 120463. <https://doi.org/10.1016/j.energy.2021.120463>
- Khan, A. y Ghosh, SK (2021). Análisis y predicción del desempeño de los estudiantes en el aprendizaje en el aula: una revisión de estudios de minería de datos educativos. *Educación y tecnologías de la información*, 26, 205–240. <https://doi.org/10.1007/s10639-020-10230-3>
- Korkmaz, C. y Correia, AP (2019). Una revisión de la investigación sobre el aprendizaje automático en la tecnología educativa. *Educativo Media International*, 56(3), 250–267. <https://doi.org/10.1080/09523987.2019.1669875>
- Kotsiantis, S., Patriarcheas, K. y Xenos, M. (2010). Un conjunto incremental combinacional de clasificadores como técnica para predecir el desempeño de los estudiantes en educación a distancia. *Sistemas basados en el conocimiento*, 23(6), 529–535. <https://doi.org/10.1016/j.knosys.2010.03.010>
- Kotsiantis, S., Pierrakeas, C. y Pintelas, P. (2004). Predecir el rendimiento de los estudiantes en el aprendizaje a distancia utilizando técnicas de aprendizaje automático. *Inteligencia artificial aplicada*, 18(5), 411–426. <https://doi.org/10.1080/08839510490442058>
- Kotsiantis, SB, Zaharakis, ID y Pintelas, PE (2006). Aprendizaje automático: una revisión de las técnicas de clasificación y combinación. *Revisión de inteligencia artificial*, 26(3), 159–190. <https://doi.org/10.1007/s10462-007-9052-3>
- Kumar, M. y Singh, AJ (2017). Evaluación de técnicas de minería de datos para la predicción del rendimiento de los alumnos. *Revista internacional de educación moderna y ciencias de la computación*, 8 (4), 25–31. <https://doi.org/10.5815/ijmecs.2017.08.04>
- McLachlan, G. (1992). Análisis discriminante y reconocimiento de patrones estadísticos. Wiley.
- Nieuwoudt, JE (2020). Investigando la asistencia a clase sincrónica y asincrónica como predictores del éxito académico en la educación en línea. *Revista de tecnología educativa de Australasia*, 36 (3), 15–25. <https://doi.org/10.14742/AJET.5137>
- Nouri, J., Saqr, M. y Fors, U. (2019). Predecir el rendimiento de los estudiantes en un aula invertida utilizando el aprendizaje automático: hacia una retroalimentación formativa automatizada basada en datos. *ICSIT 2019: décima conferencia internacional sobre sociedad y tecnologías de la información*, actas, 17(2), 79–82.
- Olson, JS y McCracken, FE (2015). ¿Vale la pena el esfuerzo? El impacto de incorporar conferencias sincrónicas en un curso por Internet. *Aprendizaje en línea*, 19(2). <https://olj.onlinelearningconsortium.org/index.php/olj/article/view/499/132>
- Peña-Ayala, A. (2014). Minería de datos educativos: una encuesta y un análisis basado en la minería de datos de trabajos recientes. *Sistemas expertos con aplicaciones*, 41(4 PARTE 1), 1432–1462. <https://doi.org/10.1016/j.eswa.2013.08.042>
- Romero, C., López, MI, Luna, JM, & Ventura, S. (2013). Predecir el desempeño final de los estudiantes a partir de la participación en foros de discusión en línea. *Informática y Educación*, 68, 458–472. <https://doi.org/10.1016/j.compedu.2013.06.009>
- Romero, C. y Ventura, S. (2013). La minería de datos en la educación. *Revisión interdisciplinaria de Wiley: minería de datos y conocimiento* Descubrimiento, 3(1), 12–27. <https://doi.org/10.1002/widm.1075>
- Romero, C. y Ventura, S. (2020). Minería de datos educativos y análisis de aprendizaje: una encuesta actualizada. *Wiley Interdisciplinario Reseñas: Minería de datos y descubrimiento de conocimientos*, 10(3), 1–21. <https://doi.org/10.1002/widm.1355>
- Romero, C., Ventura, S. y García, E. (2008). Minería de datos en sistemas de gestión de cursos: estudio de caso y tutorial de Moodle. *Informática y Educación*, 51, 368–384.
- Russell, S. y Norving, P. (2020). *Inteligencia artificial: un enfoque moderno* (4ª ed.). Prentice Hall.
- Shahri, AM, Husain, W. y Rashid, NA (2015). Una revisión sobre cómo predecir el desempeño de los estudiantes utilizando técnicas de minería de datos. *Procedia Computer Science*, 72, 414–422. <https://doi.org/10.1016/j.procs.2015.12.157>
- Siemens, G. y Baker, RSJD (2012). Analítica de aprendizaje y minería de datos educativos. *Actas de la 2ª Conferencia Internacional sobre Análisis del Aprendizaje y Conocimiento - LAK '12, Parte F1305*. doi: <https://doi.org/10.1145/2330601.2330661>
- Singh, S. y Lal, SP (2013). Evaluación de cursos educativos utilizando técnicas de aprendizaje automático. *Conferencia IEEE sobre E-Learning, e-Management y e-Services*. <https://doi.org/10.1109/IC3e.2013.6735969>
- Sisovic, S., Matetic, M. y Bakaric, MB (2016). Agrupación de datos de moodle desequilibrados para alerta temprana de fracaso estudiantil. *2016 IEEE 14º simposio internacional sobre inteligencia e informática de máquinas aplicadas (SAMI)*, 165–170. doi: <https://doi.org/10.1109/SAMI.2016.7423001>
- Mayordomo, K. (2019). Sensibilidad vs especificidad. <https://www.technologynetworks.com/analysis/articles/sensitivity-vs-specificity-318222>. Consultado el 10 de marzo de 2021.
- Tekin, A. (2014). Predicción temprana de los promedios de calificaciones de los estudiantes al graduarse: un enfoque de minería de datos. *Revista Euroasiática de Investigación Educativa*, 14(54), 207–226. <https://doi.org/10.14689/ejer.2014.54.12>
- Tomasevic, N., Gvozdenovic, N. y Vranes, S. (2020). Una descripción general y una comparación de las técnicas de minería de datos supervisadas para la predicción del rendimiento de los exámenes de los estudiantes. *Informática y Educación*, 143, 103676. <https://doi.org/10.1016/j.compedu.2019.103676>
- UNESCO. (2020). Impacto del COVID-19 en la educación. <https://en.unesco.org/covid19/educationresponse>. Consultado el 11 de febrero y 2021.
- Wang, Z., Wang, Y. y Srinivasan, RS (2018). Un novedoso enfoque de aprendizaje conjunto para ayudar a predecir el uso de energía en edificios. *ción Energía y Edificios*, 159, 109–122. <https://doi.org/10.1016/j.enbuild.2017.10.085>
- QUIÉN. (2020). Palabras de apertura del Director General de la OMS en la rueda de prensa sobre COVID-19 - 11 de marzo de 2020. <https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-marzo-2020>. Consultado el 11 de febrero de 2021.
- Yildiz, O., Bal, A. y Gulsecen, S. (2013). Modelado difuso mejorado para predecir el rendimiento académico de la distancia estudiantes de educación. *Revista internacional de investigación en educación abierta ya distancia*, 14(5), 144–165. <https://doi.org/10.19173/irrodi.v14i5.1595>

Tú, JW (2016). Identificar indicadores significativos usando datos LMS para predecir el logro del curso en el aprendizaje en línea.

Internet y Educación Superior, 29, 23–30. <https://doi.org/10.1016/j.iheduc.2015.11.003>

Zacharis, Nueva Zelanda (2015). Un enfoque multivariante para predecir los resultados de los estudiantes en cursos de aprendizaje combinado habilitados para la web. Internet y Educación Superior, 27, 44–53. <https://doi.org/10.1016/j.iheduc.2015.05.002>

Nota del editor

Springer Nature se mantiene neutral con respecto a las reclamaciones jurisdiccionales en mapas publicados y afiliaciones institucionales.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)
