


ARTÍCULO DE INVESTIGACIÓN

Acceso abierto



Prevalencia y factores predictivos del estrés percibido entre estudiantes universitarios de Bangladesh que utilizan algoritmos de aprendizaje automático

Rumana Rois1 , Manik Ray1, Atikur Rahman1 y Swapan K. Roy2

Abstracto

Antecedentes: Los problemas de salud mental relacionados con el estrés son una de las causas más comunes de sobrecarga en estudiantes universitarios a nivel mundial. Se han realizado muchos estudios para predecir la prevalencia del estrés entre los estudiantes universitarios, sin embargo, la mayoría de estos análisis se realizaron predominantemente utilizando el modelo básico de regresión logística (LR). Como alternativa, utilizamos enfoques de aprendizaje automático (ML) avanzado para detectar factores de riesgo significativos y predecir la prevalencia del estrés entre los estudiantes universitarios de Bangladesh.

Métodos: este estudio de prevalencia encuestó a 355 estudiantes de veintiocho universidades diferentes de Bangladesh utilizando preguntas sobre medidas antropométricas, académicas, estilos de vida e información relacionada con la salud, que se referían al estado de estrés percibido de los encuestados (sí o no). Se utilizó el algoritmo de Boruta para determinar los factores pronósticos significativos de la prevalencia del estrés. Los modelos de predicción se construyeron usando árbol de decisión (DT), bosque aleatorio (RF), máquina de vector de soporte (SVM) y LR, y sus desempeños se evaluaron usando parámetros de matriz de confusión, curvas de características operativas del receptor (ROC) y k- técnicas de validación cruzada de pliegues.

Resultados: Un tercio de los estudiantes universitarios reportaron estrés en los últimos 12 meses. La frecuencia del pulso de los estudiantes, las presiones arteriales sistólica y diastólica, el estado del sueño, el tabaquismo y la formación académica fueron seleccionados como características importantes para predecir la prevalencia del estrés. El rendimiento evaluado reveló que el rendimiento más alto observado de RF (exactitud = 0,8972, precisión = 0,9241, sensibilidad = 0,9250, especificidad = 0,8148, área bajo la curva ROC (AUC) = 0,8715, precisión de k-fold = 0,8983) y el más bajo de LR (exactitud=0,7476, precisión=0,8354, sensibilidad=0,8250, especificidad=0,5185, AUC=0,7822, precisión k-fold=0,7713) y SVM con núcleo polinomial de grado 2 (exactitud=0,7570, precisión=0,7975, sensibilidad=0,8630, especificidad = 0,5294, AUC = 0,7717, precisión de k veces = 0,7855). En general, el modelo de RF funciona mejor y predice el estrés de forma auténtica en comparación con otras técnicas de ML, incluidos los efectos individuales y de interacción de los predictores.

Conclusión: el marco de aprendizaje automático puede detectar los factores pronósticos significativos y predecir este problema psicológico con mayor precisión, lo que ayuda a los responsables de la formulación de políticas, las partes interesadas y las familias a comprender y prevenir esta grave crisis al mejorar las estrategias de formulación de políticas, promoción de la salud mental, y establecer servicios de asesoramiento universitario eficaces.

*Correspondencia: rois@juniv.edu

¹ Departamento de Estadística, Universidad de Jahangirnagar, Dhaka, Bangladesh
La lista completa de información del autor está disponible al final del artículo.



© El(los) autor(es) 2021. **Acceso abierto** Este artículo tiene una licencia internacional Creative Commons Attribution 4.0, que permite el uso, el intercambio, la adaptación, la distribución y la reproducción en cualquier medio o formato, siempre que se otorgue el crédito correspondiente al original autor(es) y la fuente, proporcione un enlace a la licencia Creative Commons e indique si se realizaron cambios. Las imágenes u otro material de terceros en este artículo están incluidos en la licencia Creative Commons del artículo, a menos que se indique lo contrario en una línea de crédito al material. Si el material no está incluido en la licencia Creative Commons del artículo y su uso previsto no está permitido por la regulación legal o excede el uso permitido, deberá obtener el permiso directamente del titular de los derechos de autor. Para ver una copia de esta licencia, visite <http://creativecommons.org/licenses/by/4.0/>. La renuncia a la dedicación de dominio público de Creative Commons (<http://creativecommons.org/dominiopublico/zero/1.0/>) se aplica a los datos disponibles en este artículo, a menos que se indique lo contrario en una línea de crédito a los datos.

Palabras clave: salud mental, árbol de decisión, bosque aleatorio, máquina de vectores de soporte, selección de características, matriz de confusión, ROC, validación cruzada de k-fold

Introducción

El estrés no es un diagnóstico psiquiátrico, pero está estrechamente relacionado con condiciones de salud mental que incluyen depresión, ansiedad, psicosis y trastorno de estrés postraumático [1]. El estrés puede definirse como "la incapacidad de hacer frente a una amenaza percibida (real o imaginaria) para el bienestar mental, físico, emocional y espiritual de uno que da como resultado una serie de respuestas y adaptaciones fisiológicas" [2]. Esta amenaza puede ser positiva (eustress) como graduarse o comenzar una nueva relación, o negativa, también llamada angustia, con ejemplos que incluyen la prueba académica o no poder pagar las cuotas del semestre [3]. Los estudiantes que asisten a una universidad pueden experimentar tanto eustrés como angustia en forma crónica (como roles múltiples y finanzas inadecuadas) o eventos de la vida (como reubicación y muerte) [3]. Los días universitarios de un individuo son emocional e intelectualmente más exigentes que casi cualquier otra etapa de la educación [4]. En esta etapa, un individuo enfrenta una gran cantidad de presiones y desafíos que plantean una variedad de dificultades físicas, sociales y emocionales [5].

Durante este período de transición, los estudiantes deben hacer frente a las demandas académicas y sociales que encuentran en los estudios universitarios que los ayudan en su preparación para carreras profesionales mediante la adquisición de conocimientos profesionales, habilidades transferibles y actitudes basadas en evidencia [6–9]. De acuerdo con una encuesta universitaria nacional de salud de la Asociación Nacional de Salud Mental, 1 de cada 10 estudiantes universitarios ha sido diagnosticado con depresión [10]. El último informe de la American College Health Association de 2014 indicó que aproximadamente la mitad de los estudiantes reportaron un estrés tremendo o superior al promedio en los últimos 12 meses [11]. Además, ampliar los servicios de salud mental contribuirá al logro del Objetivo de Desarrollo Sostenible (ODS) 3 de bienestar para 2030, para reducir en un tercio la mortalidad prematura por enfermedades no transmisibles a través de la prevención y el tratamiento y promover la salud mental. salud y bienestar [12].

Numerosas investigaciones se han centrado en el estudio de la prevalencia de los problemas de salud mental entre la población universitaria y los hallazgos sugieren que, en todo el mundo, un número considerable de estudiantes universitarios experimentan problemas de salud mental [4, 6–9, 13–23]. En Bangladesh, hay mucho trabajo en la literatura sobre la prevalencia de problemas de salud mental entre estudiantes universitarios y los resultados enfatizan que la prevalencia de depresión, ansiedad,

y se ha informado que el estrés llega al 54,3 %, 64,8 % y 59,0 %, respectivamente [9, 24–28].

La mayoría de los estudios relacionados con el estrés se han centrado en la predicción de la prevalencia de problemas de salud mental utilizando el modelo de regresión logística (LR). El modelado pronóstico con LR está bien establecido, particularmente para un resultado dicotómico. Aunque LR es un modelo popular de aprendizaje automático (ML) para la clasificación, estamos interesados en evaluar el rendimiento de diferentes modelos de ML, incluido LR, para predecir la prevalencia del estrés entre los estudiantes universitarios de Bangladesh. El ML en el cuidado de la salud generalmente tiene como objetivo predecir algunos resultados clínicos sobre la base de múltiples predictores [29, 30]. El potencial de ML en el cuidado de la salud es enorme, con demostraciones de herramientas basadas en ML que pueden lograr capacidades de diagnóstico y pronóstico de nivel humano o superiores que se han descrito en casi todas las especialidades clínicas [31]. El marco de ML puede explorar información más vital sobre este problema crucial de salud pública. Por lo tanto, estamos motivados para encontrar los factores de riesgo (características) y predecir la prevalencia del estrés entre los estudiantes universitarios de Bangladesh.

Materiales y métodos

Participantes y procedimientos

Realizamos un estudio transversal en línea entre estudiantes universitarios de diferentes universidades de Bangladesh de enero a marzo de 2020, justo antes del brote de COVID-19 en Bangladesh. Los participantes fueron incluidos de forma anónima y voluntaria. Los datos se recopilaron mediante un muestreo por conveniencia a través de una encuesta autoinformada en línea en las diferentes universidades del país. Considerando el nivel de significancia del 5%, margen de error aceptable del 5% ($d = 0.05$), y (En base a nuestro estudio piloto (dado que el 36,3 % de los estudiantes universitarios reportaron estrés en los últimos 12 meses en nuestro estudio piloto), el tamaño de muestra deseado se estimó siguiendo la fórmula de Cochran:

$$n_{\text{Cochran}} = \frac{Z^2 \cdot p(1-p)}{d^2}.$$

Por eso, la muestra requerida $n = 355.318$ el tamaño era ≈ 355 . Por lo tanto, se recopilaron datos de 355 participantes mediante un formulario de Google bien estructurado. Por lo tanto, no hubo cuestionarios incompletos de ningún participante. La variable objetivo, el estrés, se informó de acuerdo con su percepción del estrés con una respuesta binaria (sí = 1, no = 0). Se incluyeron variables de entrada género, año académico, sus antecedentes

(departamento) y universidad, y la actividad física relacionada con el estrés y las variables del estilo de vida, como la duración del sueño, la frecuencia del pulso (bajo = menos de 60 latidos por minuto, normal = 60 a 100 latidos por minuto, alto = más de 100 latidos por minuto) por minuto), presión arterial sistólica (PAS), presión arterial diastólica (PAD), índice de masa corporal (IMC), hábito de beber y fumar. Los estudiantes fueron clasificados según las pautas de la organización mundial de la salud como bajo peso (es decir, $IMC < 20 \text{ kg/m}^2$), peso normal (es decir, $20 \text{ kg/m}^2 < IMC < 25 \text{ kg/m}^2$), sobrepeso/

obesos (es decir, $IMC > 25 \text{ kg/m}^2$) según su valor de IMC [32]. Para la duración del sueño, se pidió a los participantes que informaran la duración promedio del sueño por día como normal (6-7 h), corta (<6 h) o larga (>7 h) [27]. Según el informe del Comité Nacional Conjunto, las categorías de presión arterial (PA) se definieron como normotensas (PA normal) si la PAS observada estaba entre 91 y 120 mmHg o la PAD estaba entre 61 y 80 mmHg; Prehipertenso si la PAS observada estaba entre 121 y 139 mmHg o la PAD estaba entre 81 y 89 mmHg, y se consideraba Hipertenso si la PAS observada era igual o superior a 140 mmHg y la PAD era igual o superior a 90 mmHg, y finalmente, Hipotensión era definida como PAS igual o inferior a 90 mmHg o PAD igual o inferior a 60 mmHg [33–35].

Cuestiones éticas

Durante todo el estudio se siguieron las directrices éticas internacionales para la investigación biomédica en seres humanos. Después de la aprobación de la propuesta de investigación, se recibió el permiso ético para la recopilación de datos del Departamento de Estadística de la Universidad de Jahangirnagar, Bangladesh. Los participantes respondieron de forma anónima a la encuesta en línea completando una carta de consentimiento informado en la primera sección del cuestionario electrónico. En el formulario de consentimiento, a todos los participantes se les proporcionó información sobre el propósito de la investigación, la confidencialidad de la información y el derecho a revocar la participación sin justificación previa.

Análisis estadístico

Este estudio tuvo como objetivo clasificar y predecir el estrés mental entre los estudiantes universitarios de Bangladesh y evaluar los factores de riesgo de su estrés utilizando diferentes modelos de clasificación ML, por ejemplo, árbol de decisión (DT), bosque aleatorio (RF), máquina de vectores de soporte (SVM) y LR. Nuestra metodología implica, en consecuencia, la recopilación de datos y el preprocesamiento, la selección de características (los factores de riesgo) utilizando el algoritmo de Boruta, la división de todo el conjunto de datos en conjuntos de datos de entrenamiento y prueba, la aplicación de modelos ML en el conjunto de datos de entrenamiento y la evaluación del rendimiento de estos modelos en el conjunto de datos de prueba y, finalmente, utilizando el modelo mejor realizado, prediga el estrés mental en función del conjunto de datos de prueba. La

el rendimiento se evaluó utilizando tres parámetros de rendimiento de la matriz de confusión, como la sensibilidad, la especificidad y la precisión, el área bajo la curva de características operativas del receptor (ROC) (AUC) y la validación cruzada k-fold. Todos los modelos de ML se realizaron utilizando el módulo scikit-learn en el lenguaje de programación Python versión 3.7.3. Solo se implementó el algoritmo Boruta para seleccionar los factores de riesgo utilizando el paquete Boruta en el lenguaje de programación R [36].

Algoritmo Boruta

Se realizó el algoritmo de Boruta para extraer los factores de riesgo relevantes para el estrés percibido de los estudiantes universitarios a partir de este conjunto de datos. Este es un algoritmo de construcción de envoltura alrededor del clasificador de RF para descubrir la relevancia y las características importantes con respecto a la variable de resultado [37]. La medida de importancia de un atributo para todos los árboles del bosque se obtiene como la pérdida de precisión de la clasificación provocada por la permutación aleatoria de valores de atributos entre objetos. De aquí en adelante, el algoritmo elimina de forma iterativa las características que una prueba estadística demuestra que son menos relevantes que las pruebas aleatorias [37].

Árbol de decisión (DT)

Un DT es una de las técnicas más simples e intuitivas en ML basado en el paradigma divide y vencerás [38].

Un DT, cuyos nodos internos son pruebas (en patrones de entrada) y cuyos nodos de hoja son categorías (de patrones), asigna un número de clase (o salida) a un patrón de entrada filtrando el patrón a través de las pruebas en el árbol [39]. Cada prueba tiene resultados mutuamente excluyentes y exhaustivos [39].

Bosque aleatorio (RF)

Un algoritmo de RF tiene hiperparámetros que especifican el número de árboles y la profundidad máxima de cada árbol (efectivamente, cuántas interacciones se consideran en el modelo), mientras que las reglas de decisión son los parámetros [40]. El RF es un enfoque de aprendizaje conjunto para la clasificación que utiliza una gran colección de DT descorrelacionados [41].

En este experimento, hemos utilizado 100 DT y Gini para el índice de impurezas para implementar el algoritmo de RF en Python.

Máquina de vectores de soporte (SVM)

Las SVM [42, 43] son métodos de aprendizaje supervisado que analizan datos y reconocen patrones. Para una tarea de aprendizaje de dos clases, un algoritmo de entrenamiento SVM construye un modelo o una función de clasificación que asigna nuevas observaciones a una de las dos clases a cada lado de un hiperplano, convirtiéndolo en un clasificador lineal binario no probabilístico. Un modelo SVM utiliza el truco del núcleo para mapear los datos en un espacio de mayor dimensión antes de resolver la tarea de ML como un problema de optimización convexo [41–

44]. Luego se predice que las nuevas observaciones pertenecen a una de las dos clases. La

lado de la partición se caen. Los vectores de soporte son los puntos de datos más cercanos al hiperplano que divide las clases [41]. Examinamos los modelos SVM utilizando el núcleo polinomial de grado 2 y el núcleo lineal para este análisis.

Regresión logística (LR)

LR es un modelo de clasificación estadística probabilística que predice la probabilidad de ocurrencia de un evento [41]. LR modela la relación entre una variable dependiente categórica y un resultado o característica categórica dicotómica. Se utiliza como modelo binario (múltiple) para predecir respuestas binarias (múltiples), el resultado de una variable dependiente categórica, en función de una o más variables independientes [38]. Este es un modelo limitado por supuestos, antes de estimar el modelo se deben cumplir todos los supuestos subyacentes, entre ellos los predictores deben ser independientes entre sí y tener una asociación significativa con la variable de resultado son los supuestos inevitables [45].

Parámetros de rendimiento de la matriz de confusión

Una matriz de confusión proporciona una representación visual de las precisiones de clase reales frente a las predichas [41]. Para visualizar el rendimiento del algoritmo de clasificación, compara la clasificación predicha con la clasificación real en forma de información de falso positivo (FP), verdadero positivo (TP), falso negativo (FN) y verdadero negativo (TN) [38, 41]. Por lo tanto, los parámetros de rendimiento son:

$$\text{Precisión} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FN} + \text{FP}}, \quad (1)$$

$$\text{Sensibilidad} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \quad (2)$$

$$\text{Especificidad} = \frac{\text{Tennese}}{\text{TN} + \text{FP}}, \quad (3)$$

$$\text{Precisión} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \quad (4)$$

donde la precisión es el número de puntos de datos clasificados correctamente por el clasificador, la sensibilidad es una medida de qué tan bien un algoritmo de clasificación clasifica los puntos de datos en la clase positiva, la especificidad es una medida de qué tan bien un algoritmo de clasificación clasifica los puntos de datos en la clase negativa, y la precisión es el número de puntos de datos clasificados correctamente de la clase positiva [38, 41].

Curva característica operativa del receptor (ROC)

Las curvas ROC ofrecen otra representación gráfica útil para los clasificadores que operan en conjuntos de datos. Fawcett [46] proporcionó una introducción completa al análisis ROC,

destacando conceptos erróneos comunes. La curva ROC muestra la sensibilidad del clasificador al graficar la tasa de verdaderos positivos con respecto a la tasa de falsos positivos. Si el clasificador es sobresaliente, la tasa de verdaderos positivos aumentará y el área bajo la curva (AUC) será cercana a 1 [38].

Validación cruzada de pliegues en K

La validación cruzada es una técnica de verificación que evalúa la capacidad de generalización de un modelo para un conjunto de datos independiente [41]. Evalúa el desempeño de varias funciones de predicción. En la validación cruzada k-fold, el conjunto de datos de entrenamiento se divide arbitrariamente en k submuestras (o pliegues) mutuamente excluyentes de igual tamaño. El modelo se entrena k veces (o pliegues), donde cada iteración usa uno de los k submuestras para pruebas (validación cruzada), y las submuestras restantes (k-1) se aplican para entrenar el modelo. Los k resultados de la validación cruzada se promedian para estimar la precisión como una única estimación [41]. Para este pequeño tamaño de muestra, aplicamos técnicas de validación cruzada de tres, cinco y diez veces para evaluar el desempeño de los clasificadores.

Resultados

En esta encuesta han participado un total de 355 estudiantes de 28 universidades diferentes de Bangladesh, con la mayor proporción de respuestas de la Universidad Jahangirnagar (56,1%), seguida de la Universidad de Dhaka (5,9%) y la Universidad de Rajshahi (5,6%). La información detallada se encuentra en el archivo complementario. Entre los participantes, 204 eran mujeres (57,5%), 22,5% tenían sobrepeso/obesidad, 15,8% eran fumadores de cigarrillos, 8,5% eran alcohólicos y 30,7% de los estudiantes universitarios reportaron estrés en los últimos 12 meses. La mayoría de los estudiantes tenían pulso normal (76,9%), 63,4% dormían normalmente, 77,5% tenían PA normotensa para PAS y 76,9% PA normotensa para PAD (Tabla 1). Poco más de la mitad de la muestra total, el 62,3% (n=221) eran estudiantes de posgrado, seguidos por el 13% (n=46) estudiantes universitarios de primer año. La muestra incluyó 37,7% (n=134) de estudiantes de posgrado, de los cuales 33,6% (n=45) reportaron estrés. Los estudiantes de posgrado tenían menos probabilidades de estar estresados que los estudiantes de pregrado, ya que el 29,0% (n=64) de los estudiantes de posgrado informaron estrés. La proporción más alta de participantes 51,5% (n=183) eran de formación científica, seguida por 18,3% (n=65) de artes. Los estudiantes estresados eran más propensos a ser hombres (35,1%), estudiantes de medicina (40%), estudiantes de primer año de pregrado (41,3%), no fumadores de cigarrillos (39,3%), pulso bajo (96,5%), durmientes normales (34,7%), con sobrepeso/obesidad (36,3%), tenían PAS hipotensa (100%) o hipertensa (100%) y PAD hipotensa (100%), como se muestra en la Tabla 1.

Tabla 1 Distribución de frecuencias y relación con el estrés en estudiantes universitarios

Variables	Total 355 Estrés (n=109; 30,7%)			
	norte (%)	Sí (%)	χ ²	valor p
Género				
Mujer	204 (57,5)	56 (27,5)	2.386	0.131
Masculino	151 (42,5)	53 (35,1)		
Universidad				
1. Universidad de Jahangirnagar	199 (56,1)	59 (29,6)	38.811	0.066
2. Universidad de Daca	21 (5,9)	5 (23,8)		
...		
27. Universidad Nacional	2 (0,6)	1 (50,0)		
28. Universidad de Rajshahi	20 (5,6)	7 (35,0)		
Antecedentes				
Letras	65 (18,3)	20 (30,8)	2.891	0.576
Ciencias	183 (51,5)	50 (27,3)		
Comercio	40 (11,3)	14 (35,0)		
Médico	30 (8,5)	12 (40,0)		
Ingeniería	37 (10,4)	13 (35,1)		
Año académico				
1er año	46 (13,0)	19 (41,3)	3.506	0.477
2do año	33 (9,3)	8 (24,2)		
3er año	31 (8,7)	10 (32,3)		
4to año	24 (6,8)	8 (33,3)		
Maestros	221 (62,3)	64 (29,0)		
La frecuencia del pulso				
Bajo	57 (16,1)	55 (96,5)	200,75	<0,001*
Normal	273 (76,9)	32 (11,7)		
Alto	25 (7,0)	22 (88,0)		
Alcohólico				
sí	30 (8,5)	9 (30,0)	0,008	0,930
No	325 (91,5)	100 (30,8)		
Tabaquismo				
sí	56 (15,8)	22 (29,1)	2.301	0.129
No	299 (84,2)	22 (39,3)		
Hora de dormir				
Menos de lo normal	29 (8,2)	9 (31)	5.441	0.066
Normal	225 (63,4)	78 (34,7)		
Mas de lo normal	101 (28,5)	22 (21,8)		
PAS				
Hipotensión	19 (5,4)	19 (100)	84.320	<0.001*
normotenso	275 (77,5)	59 (21,5)		
prehipertenso	48 (13,5)	18 (37,5)		
hipertenso	13 (3,7)	13 (100)		
PAD				
Hipotensión	13 (3,7)	13 (100)	79,554	<0,001*
normotenso	273 (76,9)	63 (23,1)		
prehipertenso	45 (12,7)	11 (24,4)		
hipertenso	24 (6,8)	22 (91,7)		
IMC				
bajo peso	77 (21,7)	24 (31,2)	1.710	0.425
Peso normal	198 (55,8)	56 (28,3)		
Sobrepeso/obesidad	80 (22,5)	29 (36,3)		

Tabla 1 (continuación)

*Estadísticamente significativo al nivel 0.05

La Tabla 1 también muestra que los participantes estresados fueron significativamente más propensos que los participantes no estresados a tener un pulso bajo ($\chi^2=200,75$, valor de $p<0,05$), hipotensión o PAS hipertensa ($\chi^2=84,320$, valor de $p<0,05$).), y presentaba PAD hipotensa ($\chi^2=79,554$, valor de $p<0,05$).

Selección de características

La Figura 1 revela que, con la ayuda del algoritmo de Boruta, se seleccionaron seis variables (frecuencia del pulso, PAS, PAD, estado del sueño, tabaquismo, antecedentes [departamento]) entre diez variables encuestadas como factores de riesgo para predecir el estrés entre los estudiantes universitarios de Bangladesh. . La frecuencia del pulso, el estado del sueño, la PAS y la PAD de los estudiantes fueron las características confirmadas y el hábito de fumar y los antecedentes fueron las características tentativas para clasificar su estrés mental. De aquí en adelante, estas seis variables se utilizaron para evaluar el rendimiento de los algoritmos de ML.

Evaluación de modelos de aprendizaje automático

El rendimiento de los modelos ML como DT, RF, SVM y LR se evaluó utilizando cuatro parámetros de rendimiento de la matriz de confusión (Tabla 2), el área bajo la curva ROC (Fig. 2) y la validación cruzada de k-fold enfoques (Tabla 3). Considerando el 70 % de las observaciones como datos de entrenamiento y el 30 % de las observaciones como datos de prueba con las semillas aleatorias 2370–2380 para once ejecuciones diferentes usando el módulo scikit-learn, estimamos el puntaje promedio de exactitud, sensibilidad, especificidad y precisión de DT , RF, SVM y LR para predecir el estrés entre los estudiantes universitarios y los resultados se ilustran en la Tabla 2. La Tabla 2 también

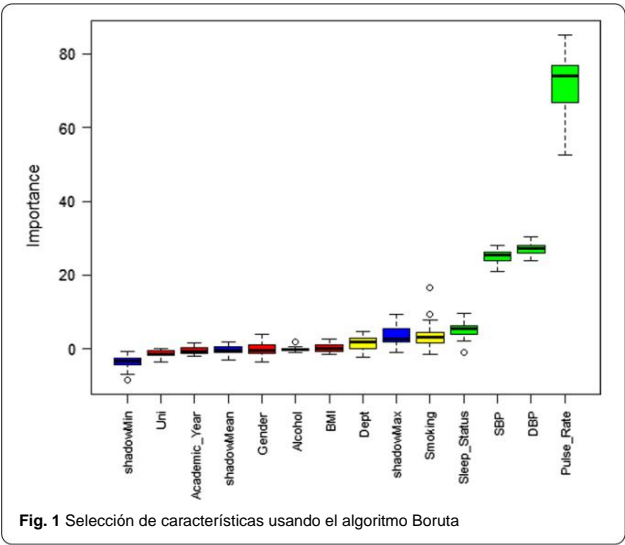
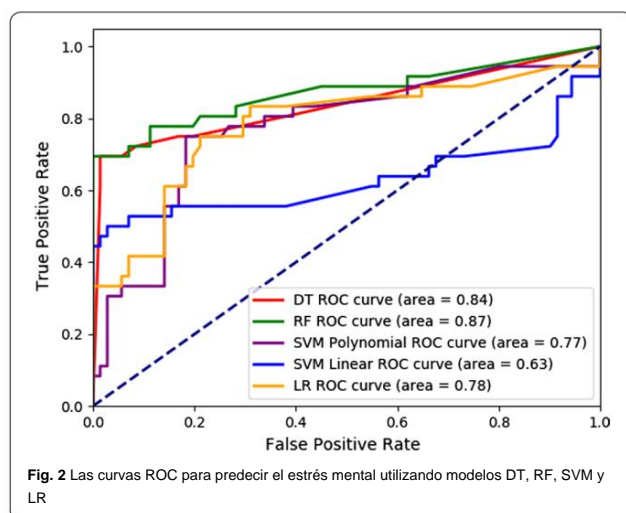


Fig. 1 Selección de características usando el algoritmo Boruta

Tabla 2 Exactitud, sensibilidad, especificidad y precisión de diferentes modelos de ML

Modelos	Exactitud		Sensibilidad		especificidad		Precisión	
	Significar	SE	Significar	SE	Significar	SE	Significar	SE
DT	0.8845	0.017	0.8908	0.027	0.8639	0.076	0.9581	0.024
RF	0.8929	0.014	0.8953	0.027	0.8853	0.075	0.9653	0.021
SVM (núcleo polinomial)	0.7782	0.035	0.8504	0.047	0.6047	0.065	0.8406	0.039
SVM (núcleo lineal)	0.8054	0.039	0.8460	0.045	0.7188	0.172	0.8969	0.068
LR	0.7723	0.037	0.8160	0.045	0.6175	0.094	0.8848	0.031

Media Media de las diferentes puntuaciones de las ejecuciones repetidas, SE Error estándar de las diferentes puntuaciones de las ejecuciones repetidas



muestra las estimaciones de incertidumbre del parámetro utilizando el error estándar (SE) de estos parámetros de rendimiento estimados, el error estándar es la desviación estándar de estas estimaciones. El puntaje promedio estimado más alto de los parámetros de rendimiento y el SE más bajo de ellos se indican en negrita en las Tablas 2 y 3, un valor en negrita indica un mejor rendimiento del modelo ML correspondiente. Los desempeños evaluados revelaron que el modelo RF fue el más eficiente para predecir el estrés entre todos los modelos ML examinados en base al valor más alto de la

parámetros de rendimiento estimados y con el menor valor de incertidumbre de dichas estimaciones en todos los casos. Por ejemplo, el modelo RF proporcionó el 89,3 % de predicciones precisas (es decir, precisión=0,8929) con SE=0,014, el 89,5 % de casos positivos que se predijeron como positivos (es decir, sensibilidad=0,8953) con SE=0,027, el 88,5 % de casos negativos que se pronosticaron como negativos (es decir, especificidad=0,8853) con SE=0,075, y 96,5% de predicciones positivas que fueron correctas (es decir, precisión=0,9653) con SE=0,021.

La Figura 2 ilustra el AUC estimado de los modelos DT, RF, SVM y LR, que se ejecutaron con el módulo scikit-learn en Python 3.7.3 al considerar el 70 % de las observaciones como datos de entrenamiento y el 30 % de las observaciones como datos de prueba con el resultado aleatorio. semilla 1439. Para predecir la prevalencia de estrés mental en los últimos 12 meses entre estudiantes universitarios, el AUC estimado fue 0.8388, 0.8715, 0.7717, 0.6285 y 0.7822 utilizando los modelos ML DT, RF, SVM con el núcleo polinomial de grado 2, SVM con kernel lineal y LR, respectivamente. El algoritmo de RF se desempeñó mejor con el AUC máximo entre todos los modelos de ML examinados. La validación cruzada de K veces se realizó para repeticiones triples, quintuples y de 10 veces con la semilla aleatoria 1 y el argumento shufe 'Verdadero', y los resultados se organizan en la Tabla 3. El modelo RF funcionó mejor en tres veces, quintuples y 10- Doble las validaciones cruzadas basadas en los puntajes de precisión más altos, es decir, 88.4%, 89.3% y 89.8%, respectivamente, y en general la incertidumbre más ba

Tabla 3 Resultado de la validación cruzada K-Fold de modelos ML

Modelos	Triple		Cinco veces		10 veces	
	MACc	SE	MACc	SE	MACc	SE
DT	0.8759	0.0419	0.8901	0.0138	0.8870	0.0361
RF	0.8844	0.0291	0.8929	0.0126	0.8983	0.0338
SVM (núcleo polinomial)	0.7718	0.0215	0.7915	0.0559	0.7855	0.0661
SVM (núcleo lineal)	0.8085	0.0072	0.8338	0.0187	0.8309	0.0383
LR	0.7830	0.0396	0.7718	0.0566	0.7713	0.0669

Puntajes MACc Mean of Accuracy de cada pliegue, puntajes SE Standard Error of Accuracy

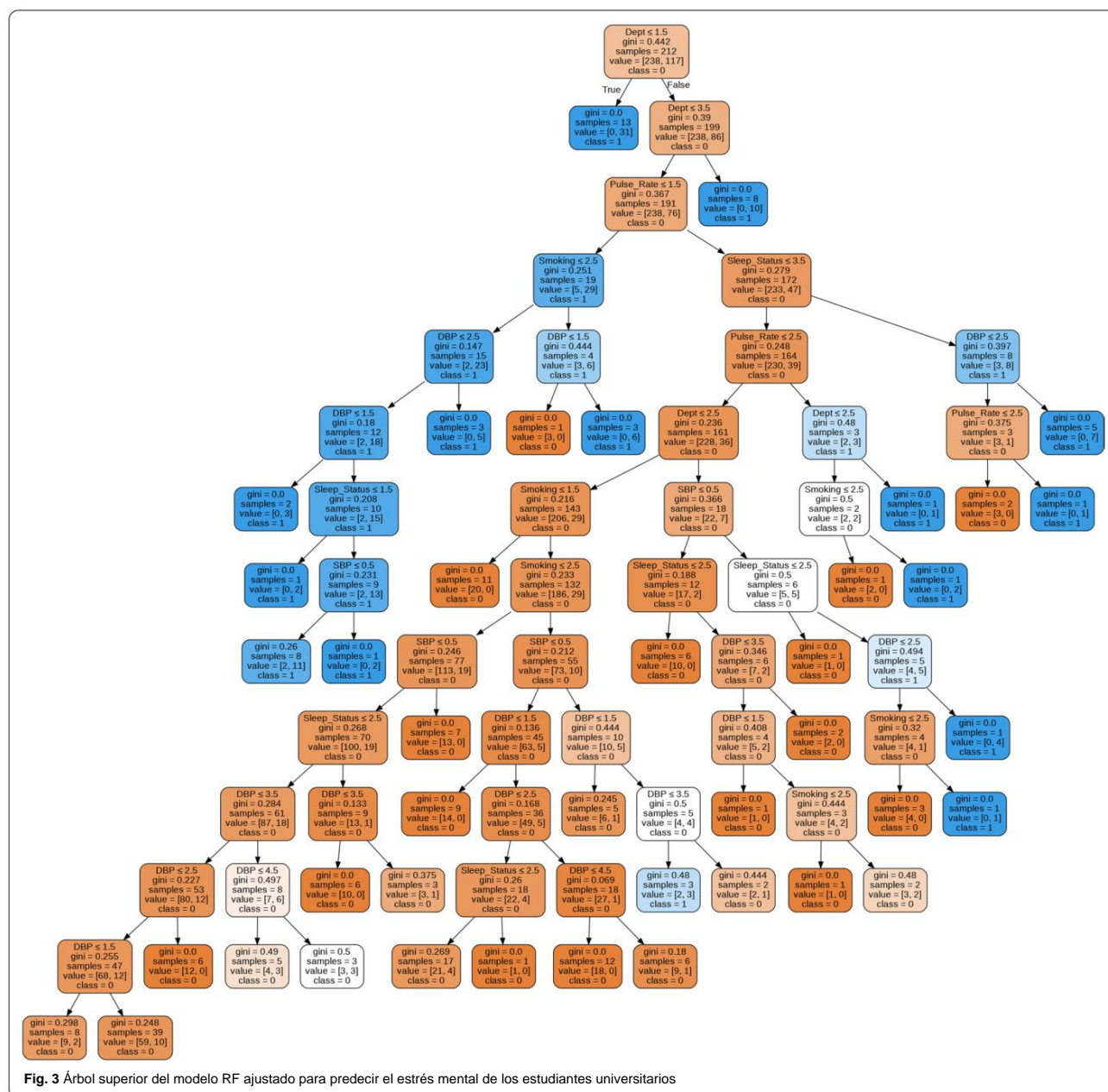
las estimaciones de los parámetros, es decir, 0,0291, 0,0126 y 0,0338, respectivamente, como se muestra en la Tabla 3.

Para predecir el estrés mental en los últimos 12 meses entre los estudiantes universitarios de Bangladesh, el algoritmo RF funcionó mejor que los algoritmos DT, SVM y LR según la medida de precisión, el ROC y los enfoques de validación cruzada de k-fold.

Modelo para predecir el estrés

Para todo el conjunto de datos, por lo tanto, el modelo ML mejor realizado, el modelo RF, se ajustó para predecir el estrés usando

los factores significativos seleccionados: frecuencia del pulso, PAS, PAD, estado del sueño, hábito de fumar y Antecedentes (departamento) de los estudiantes, y el árbol superior del bosque se visualiza en la Fig. 3. Todos los nodos tienen cinco partes (fea pregunta de la naturaleza, gini, muestras, valor y clase) con una pregunta basada en el valor de una característica, excepto que los nodos de la hoja terminal tienen cuatro partes (gini, muestras, valor y clase) [47]. La parte 'gini' indica la impureza de Gini del nodo, que es la impureza de Gini ponderada promedio que disminuye a medida que la ruta se mueve hacia abajo en el árbol, 'muestras' es el número de observaciones en el nodo, 'valor' es el



número de muestras en cada clase, y 'clase' indica la clasificación mayoritaria para puntos en el nodo ('clase' es la predicción para todas las muestras en el nodo hoja) [47].

La pregunta de cada función tiene una respuesta Verdadero (nodos de la izquierda) o Falso (nodos de la derecha) que divide el nodo. Según la respuesta a la pregunta, un punto de datos se mueve hacia abajo en el árbol y llega a un nodo de hoja (la decisión final). Además, la hoja de color azul indica una predicción sobre estudiantes estresados y la hoja de color naranja indica una predicción sobre estudiantes no estresados, como se muestra en la Fig. 3. Para predecir los datos de cualquier estudiante, simplemente muévase hacia abajo en el árbol. en la Fig. 3, usando la respuesta a la pregunta de la característica hasta llegar a un nodo hoja donde la clase es la predicción.

La Tabla 4 organiza esta ruta de decisión para los datos proporcionados por cinco estudiantes sobre su frecuencia del pulso, PAS, PAD, hábito de fumar, estado de sueño y antecedentes (Depto) para predecir su condición de estrés utilizando el modelo RF ajustado (Fig. 3).

El análisis de LR reveló además que los estudiantes fumadores de cigarrillos tenían 4,112 veces más probabilidades (OR=4,112, intervalo de confianza (IC) del 95 % 1,591–10,628, valor de $p<0,05$) de estar estresados que los no fumadores. Los encuestados que tenían una frecuencia de pulso normal tenían menos probabilidades (OR = 0,002, IC del 95 % 0,000–0,013, valor de $p<0,05$), y que tenían una frecuencia de pulso superior a la normal eran menos probables (OR = 0,037, IC del 95 %: 0,004–0,389, valor de $p<0,05$) para estar más estresados que aquellos que tenían un pulso bajo (Tabla 5).

El modelo LR ajustado en la Tabla 5 ilustra que los estudiantes el tabaquismo y la frecuencia del pulso fueron solo los dos factores significativos para estimar la prevalencia del estrés y los intervalos de confianza (IC) indeterminados para otros dos predictores, es decir, PAS (IC del 95 % = 0,000, –) y PAD (IC del 95 % = 0,000). , –). Sin embargo, el tabaquismo no se puede considerar al ajustar el modelo LR, ya que este factor no tiene una asociación significativa con la variable de resultado (Tabla 1). La prueba de chi-cuadrado en la Tabla 1 reveló que la frecuencia del pulso, la PAS y la PAD fueron solo los tres factores significativos para el estrés de los estudiantes. Además, la frecuencia del pulso, la PAS y la PAD de los estudiantes se asociaron significativamente entre sí, por ejemplo, la frecuencia del pulso se asoció significativamente con la PAS ($\chi^2=230,663$, valor de $p<0,01$) y la PAD ($\chi^2=247,583$, valor de $p<0,01$).), y la asociación entre la PAS y la PAD de los estudiantes fue ($\chi^2=415,105$, valor $p<0,01$)

Tabla 5 Odds ratios (OR) con IC del 95% y valores de p obtenidos del modelo LR

Variables	O	(95% IC)	valor p
La frecuencia del pulso			
Bajo (ref.)	1.000	–	–
Normal	0.002	(0,000–0,013)	<0.001*
	0.037	(0,004–0,389)	0.006*
Tabaquismo alto			
No (ref.)	1.000	–	–
sí	4.112	(1.591–10.628)	0.004*
Hora de dormir			
Menos de lo normal (ref.)	1.000	–	–
Normal	5.244	(0.811–33.911)	0.082
Mas de lo normal	5.660	(0,808–39,650)	0.081
PAS			
Hipotensión (ref.)	1.000	–	–
normotenso	0.000	(0.000–)	0.998
prehipertenso	0.000	(0.000–)	0.998
hipertenso	0.315	(0.000–)	1.000
PAD			
Hipotensión (ref.)	1.000	–	–
normotenso	0.000	(0.000–)	0.998
Antecedentes	0.000	(0.000–)	0.999
Hipertensivos	0.000	(0.000–)	0.999
Prehipertensos			
Artes (ref.)	1.000	–	–
Ciencias	1.428	(0,456–4,474)	0.541
Comercio	0.655	(0,117–3,682)	0.631
Médico	2.925	(0,545–15,702)	0.211
Ingeniería	3.210	(0,814–12,665)	0.096

OR=1 para la categoría de referencia

*Significativo al nivel del 5%

también significativo. En consecuencia, solo se necesita usar una variable entre los tres factores significativos, es decir, la frecuencia del pulso, la PAS y la PAD de los estudiantes, para ajustar correctamente el modelo LR en este análisis.

Tabla 4 Predicción del estrés de estudiantes universitarios utilizando el modelo RF ajustado

La frecuencia del pulso	PAS	PAD	De fumar	departamento	Estado de sueño	Estrés previsto
Alto	hipertenso	hipertenso	No	Letras	Normal	Estresado
Normal	Hipotensión	Hipotensión	No	Ciencias	Mas de lo normal	no estresado
Alto	normotenso	Hipotensión	No	Médico	Normal	no estresado
Normal	prehipertenso	prehipertenso	sí	Ingeniería	Menos de lo normal	Estresado
Bajo	normotenso	Hipotensión	sí	Médico	Menos de lo normal	Estresado

Discusión

Los estudiantes universitarios son más vulnerables al estrés y otros problemas de salud mental, lo que puede afectar negativamente su salud y rendimiento académico [48–50]. La prevalencia mundial de niveles moderados a extremadamente graves es del 60,8 % para la depresión, el 73 % para la ansiedad y el 62,4 % para el estrés [6–8, 18–22]. Como resultado, la preocupación pública por la salud mental de los estudiantes universitarios ha ido en aumento y su estrés se ha convertido en un concepto notorio en la salud pública. Motivado por una preocupación de salud pública tan notable, esta investigación realizó un estudio de prevalencia para encontrar los factores significativos y la predicción del estrés entre estudiantes universitarios en Bangladesh utilizando diferentes modelos de ML. Este estudio de prevalencia mostró que un tercio de los estudiantes universitarios reportaron estrés en los últimos 12 meses.

Los resultados del estudio revelan que la frecuencia del pulso, la PAS, la PAD, el estado de sueño, el estado de tabaquismo y los antecedentes de los estudiantes universitarios fueron los principales factores significativos para su estrés utilizando el algoritmo de selección de características de ML: Boruta. Sin embargo, la frecuencia del pulso, la PAS y la PAD de los estudiantes fueron solo los factores significativos de su estrés usando la prueba convencional de chi-cuadrado. Los estudiantes estresados eran más propensos a ser estudiantes de medicina (dos quintos), no fumadores de cigarrillos (menos de dos quintos), durmientes normales (más de un tercio), pulso bajo (menos de un entero), hipotensión (PAS exactamente un entero) o PAS hipertensa (exactamente un todo) y tenía PAD hipotensa (exactamente un todo). Existen fuertes diferencias de estrés y salud mental entre los estudiantes de pregrado y posgrado [51], el año académico no fue un factor significativo para nuestro estudio. Observamos que alrededor de dos quintas partes de los estudiantes de primer año, seguidos de más de un tercio de los estudiantes de pregrado de cuarto año, estaban estresados, mientras que más de dos séptimas partes de los estudiantes de posgrado estaban estresados.

El género fue un factor insignificante para la predicción del estrés, menos de dos séptimas partes de las estudiantes mujeres y más de un tercio de los estudiantes varones fueron percibidas como estresadas en el último año. Estos hallazgos de la investigación actual también difieren de los estudios anteriores [4, 49, 52–54].

Evaluamos el rendimiento de los modelos ML como DT, RF, SVM y LR para predecir el estrés de los estudiantes universitarios utilizando cuatro parámetros de rendimiento de la matriz de confusión, el AUC y los enfoques de validación cruzada de k-fold. El modelo de RF se realizó mejor para predecir el estrés en todas las situaciones utilizando once ejecuciones repetidas con las estimaciones medias más altas de los parámetros de rendimiento y, en general, las estimaciones de incertidumbre más bajas de esos parámetros, es decir, 89,3 % de precisión, 96,5 % de precisión, 89,5 % de sensibilidad, 88,5 % de especificidad, 87,2 % de AUC, más del 88 % de precisión en todas las técnicas de validación cruzada de 3, 5 y 10 pliegues. El modelo RF se consideró el efecto individual y de interacción de todos los

factores seleccionados para predecir el estrés percibido de los estudiantes universitarios. Siguiendo la ruta de la Fig. 3, para cualquier estudiante individual con los datos dados, su estrés percibido se puede predecir como se muestra en la Tabla 4. Por otro lado, el modelo LR no pudo estimar el intervalo de confianza para los dos predictores significativos (SBP y DBP) e ilustró significativamente solo dos predictores, es decir, el estado de fumador de los estudiantes (que no tiene una asociación significativa con el estrés) y la frecuencia del pulso. Esta salida incompleta se observa debido a una estimación inadecuada del modelo LR. Como el modelo LR requiere cumplir con todos los supuestos subyacentes antes de estimar el modelo, entre ellos, los predictores que tienen una asociación significativa con la variable de resultado y su independencia (para evitar el problema de multicolinealidad) son los principales supuestos que deben cumplirse. En este análisis, solo una variable entre la frecuencia del pulso de los estudiantes, la PAS y la PAD se usará como variable predictora para estimar correctamente el modelo LR, ya que estas variables se asociaron significativamente (usando la prueba de chi-cuadrado en la Tabla 1) con estrés y tenían una asociación significativa entre ellos. Por lo tanto, para superar el problema de multicolinealidad, solo una variable debe involucrarse en la estimación del modelo LR, de lo contrario, los resultados serán engañosos. Además, el modelo RF no requiere ninguna suposición al estimar el modelo. Por lo tanto, considerando el mejor desempeño, el modelo RF será mejor y más auténtico (en términos de cumplir con los supuestos) para predecir el estrés percibido de los estudiantes universitarios en este estudio.

Los estudios también han demostrado que los problemas de salud mental entre los estudiantes universitarios están aumentando tanto en número como en gravedad [55]. Los problemas de salud mental pueden ser una gran fuente de sufrimiento psicológico y aumentar el riesgo de conductas suicidas [6, 8, 18, 20, 56, 57]. Por lo tanto, es vital comprender y luego ofrecer un apoyo aceptable, efectivo y accesible para este grupo potencialmente vulnerable [58]. La consejería es la intervención que se ofrece de manera más consistente y se han demostrado resultados positivos en los servicios que ofrecen terapia psicodinámica, terapia breve estructurada y terapia integradora [59, 60].

Los servicios de asesoramiento universitario en Australia, el Reino Unido y los EE. UU. informan aumentos en la búsqueda de ayuda, con más estudiantes que presentan problemas más graves [49, 61–63]. Aunque no existe una conciencia notable sobre los servicios de asesoramiento universitario en Bangladesh, se llevó a cabo un número razonable de investigaciones para abordar la prevalencia de problemas de salud mental entre estudiantes universitarios [9, 24–28, 64, 65], incluso durante la pandemia de COVID-19. pandemia [66–68].

Los estudios anteriores han informado que la prevalencia del estrés entre los estudiantes universitarios de Bangladesh es tan alta como tres quintos del total de encuestados [9, 24–28]. Sin embargo, nuestros hallazgos revelaron que un tercio de los estudiantes universitarios

los estudiantes reportaron estrés en los últimos 12 meses. Esta menor tasa de prevalencia de estrés se observó cuando se informó a los estudiantes sobre su sentimiento de estrés de los últimos 12 meses mediante una pregunta de respuesta binaria (Sí o No), que es una de las principales limitaciones de este estudio. Además, otras limitaciones importantes son el pequeño tamaño de la muestra para este tipo de análisis y el uso de una muestra de conveniencia, por lo que los estudiantes de la encuesta pueden no ser representativos de la población general de estudiantes de Bangladesh. En lugar de utilizar un patrón de respuesta binaria, cualquier escala estructurada como la Escala de Estrés Percibido (PSS) o la escala de depresión, ansiedad y estrés (DASS-21) con muestras más grandes y más representativas, y utilizando el marco ML puede ser más informativo para estimar la prevalencia del estrés de los estudiantes universitarios en Bangladesh.

A pesar de las limitaciones del estudio, creemos que nuestro estudio tiene varias ventajas atractivas en la investigación de salud pública universitario y la política de Bangladesh para apoyar a este grupo potencialmente vulnerable. La prueba de chi-cuadrado convencional identificó solo tres variables (frecuencia del pulso, PAS y PAD) como factores significativos que probablemente sean el resultado del estado de estrés del estudiante, mientras que el marco ML identificó seis variables (frecuencia del pulso, PAS, PAD), estado del sueño, tabaquismo y antecedentes) como factores significativos para predecir el estrés en este análisis. No hace falta decir que este estudio introduce la aplicación de diferentes modelos ML en la predicción del estrés de los estudiantes universitarios, por ejemplo, DT y RF, que no requieren suposiciones y son muy fáciles (disponibles) de implementar en cualquier software estándar. Mientras que el clasificador popular LR requiere cumplir con todos los supuestos subyacentes antes de estimar el modelo, entre ellos, los predictores deben ser independientes entre sí y tener una asociación significativa con la variable de resultado son los supuestos inevitables. Por lo tanto, este modelo de pronóstico de uso común es difícil de estimar correctamente y una estimación incorrecta puede dar como resultado información engañosa. Los investigadores pueden darse cuenta de las limitaciones del popular modelo LR por su función de suposiciones limitadas a partir de los resultados de nuestro estudio. Para implementar el modelo LR de manera auténtica para este estudio, solo se necesita usar una variable de la frecuencia del pulso, la PAS y la PAD de los estudiantes para predecir el estrés de los estudiantes universitarios; luego, los resultados estimados del modelo serán correctos pero menos informativos.

Además, el modelo RF incluyó todas estas seis variables significativas para predecir el estrés utilizando sus efectos individuales y de interacción. Entre estos seis factores significativos, la frecuencia del pulso, la PAS y la PAD del estudiante son las consecuencias físicas de su estrés, el tabaquismo es una estrategia negativa para afrontar el estrés y los antecedentes son una causa de su estrés. Por lo tanto, nuestro estudio, aunque basado en una muestra pequeña, encuentra que la presión de estudio de los estudiantes universitarios de Bangladesh tiene consecuencias notables en su salud física y mental y desarrollar estrés negativo.

estrategias. El asesoramiento para estudiantes universitarios puede ayudar a los estudiantes a identificar los problemas emocionales causados por el estrés del estudio y explicar por qué las cosas se salen de control. La consejería estudiantil también puede proteger a un estudiante de las estrategias negativas comunes de afrontamiento del estrés al ayudarlos a notar los signos de esos métodos de afrontamiento inútiles temprano y romper los hábitos dañinos antes de que tomen el control de su vida. Teniendo en cuenta la alta precisión en la predicción, el mejor rendimiento y la característica libre de suposiciones, el modelo RF será más auténtico e informativo utilizando la muestra grande representativa del país con un cuestionario detallado para predecir el estrés percibido de los estudiantes universitarios. Hay una conciencia cada vez mayor de la investigación para abordar el riesgo elevado de problemas de salud mental en estudiantes universitarios en Bangladesh, pero existe una grave escasez del sistema de salud, los servicios de asesoramiento potencialmente vulnerable.

Conclusión

Este estudio proporciona más evidencia del hallazgo de una elevada tasa de prevalencia de estrés entre los estudiantes universitarios de Bangladesh. Este problema psicológico es muy amenazante ya que puede afectar la salud, el rendimiento académico y la capacidad de los estudiantes para desarrollar sus carreras profesionales. Además, es necesario detectar y comprender la magnitud de este problema y, por lo tanto, habilitar intervenciones adecuadas y apropiadas para este grupo vulnerable.

El marco de ML puede detectar los factores pronósticos significativos y predecir este problema psicológico con mayor precisión, lo que ayuda a los responsables de la formulación de políticas, las partes interesadas y las familias a comprender y prevenir esta grave crisis al mejorar las estrategias de formulación de políticas, la promoción de la salud mental y el establecimiento de servicios efectivos de consejería universitaria.

Agradecimientos

Los autores agradecen a todos los participantes de 28 estudiantes universitarios por dar su tiempo voluntariamente en la recolección de datos durante la encuesta. Nos gustaría agradecer al editor técnico ya los revisores por las discusiones constructivas que llevaron a la mejora del documento.

Contribuciones de los autores

RR y AR analizaron conjuntamente los datos, redactaron y revisaron el manuscrito. RR concebido y supervisado el estudio. MR recopiló datos y realizó el análisis estadístico inicial. SKR revisó críticamente y editó el manuscrito. Todos los autores leyeron y aprobaron el manuscrito final.

Fondos

Los autores no recibieron financiación específica para este trabajo.

Disponibilidad de datos y materiales.

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

Declaraciones

Aprobación ética y consentimiento para participar

Se recopilaron datos primarios y los participantes no recibieron ningún beneficio económico, y se mantuvo el anonimato para garantizar la confidencialidad y confiabilidad de los datos. Este estudio se realizó en línea en total conformidad con las pautas éticas internacionales para la investigación biomédica en la investigación con participantes humanos.

Consentimiento para publicación

Todos los participantes dieron su consentimiento informado antes de participar en la encuesta. También brindan su consentimiento para publicar los resultados analíticos de esta encuesta sin su información identificable.

Interés en competencia

Los autores declaran no tener ningún conflicto de intereses.

Detalles del autor

- ¹ Departamento de Estadística, Universidad de Jahangirnagar, Dhaka, Bangladesh.
- ² Fundación de Lactancia Materna de Bangladesh (BBF), Instituto de Salud Pública, Dhaka, Bangladesh.

Recibido: 12 abril 2021 Aceptado: 3 noviembre 2021
Published online: 27 November 2021

Referencias

1. El estrés y nuestra salud mental: ¿cuál es el impacto y cómo podemos abordarlo? MQ Mental Health 2018. <https://www.mqmentalhealth.org/stress-and-mental-health> . Consultado el 16 de mayo de 2018.

2. Seaward BL. Manejo del estrés: Principios y estrategias para la salud y el bienestar (3.ª ed.) 2002. Boston, MA: Jones and Bartlett Publishers.

3. Oswalt SB, Riddock CC. Qué hacer ante el agobio: Estudiantes de posgrado, estrés y servicios universitarios. Coll Stud Af J. 2007;27(1):24–44.

4. Saleem S, Mahmood Z. Problemas de salud mental en estudiantes universitarios: un estudio de prevalencia. FWU J Ciencias Sociales. 2013;7(2):124–30.

5. Rodgers LS, Tennison LR. Una evaluación preliminar del trastorno de adaptación entre estudiantes universitarios de primer año. Arch Psychiatr Nurs. 2009;23(3):220–30.

6. Bayram N, Bilgel N. Prevalencia y correlaciones sociodemográficas de depresión, ansiedad y estrés entre un grupo de estudiantes universitarios. Soc Psiquiatría Psiquiatra Epidemiol. 2008, 43 (8): 667–72.

7. Kulsoom B, Afsar NA. Estrés, ansiedad y depresión entre estudiantes de medicina en un entorno multiétnico. Neuropsychiatr Dis Treat. 2015;11:1713–22.

8. Haq UL, Dar MA, Aslam IS, Mahmood QK. Estudio psicométrico de la depresión, la ansiedad y el estrés en estudiantes universitarios. J Salud Pública. 2018;26(2):211–7.

9. Mamun MA, Hossain MS, Griffiths MD. Problemas de salud mental y predictores asociados entre estudiantes de Bangladesh. Adicto a la salud Int J Ment. 2019;1–15.

10. Asociación Nacional de Salud Mental. Encontrando esperanza y ayuda: Iniciativa piloto para estudiantes universitarios y depresión, 2006. <http://www.nmha.org/camh/college/index.cfm>.

11. Asociación Americana de Salud Universitaria. American College Health Association-National College Health Assessment II: Resumen ejecutivo del grupo de referencia Primavera de 2014. <https://www.acha.org>. Consultado el 6 de marzo de 2019.

12. Organización Mundial de la Salud. Invertir en el tratamiento de la depresión y la ansiedad conduce a un retorno cuádruple, 2016. <https://www.who.int/news/item/13-04-2016-invertir-en-tratamiento-para-la-depresión-y-la-ansiedad-conduce-a-un-retorno-cuádruple>.

13. Adewuya AO. Prevalencia del trastorno depresivo mayor en Nigeria estudiantes universitarios con problemas relacionados con el alcohol. Psiquiatría Gen Hosp. 2006; 28:169–73.

14. Nordin NM, Talib MA, Yaacob SN. Personalidad, Soledad y Mental Salud entre los estudiantes universitarios de las universidades de Malasia. Eur J Sci Res. 2009;36(2):285–98.

15. Ovuga E, Boardman J, Wasserman D. Estudiante de pregrado mental salud en la Universidad de Makerere, Uganda. Psiquiatría Mundial. 2006;5(1):51–2.

16. Verger P, Guagliardo V, Gilbert F, Rouillon F, et al. Trastornos psiquiátricos en estudiantes de seis universidades francesas: prevalencia de 12 meses, comorbilidad, deterioro y búsqueda de ayuda. Soc Psiquiatría Epidemiol Psiquiátrico. 2009;45(2):189–99.

17. Seim RW, Spates CR. La prevalencia y comorbilidad de las fobias específicas en estudiantes universitarios y su interés en recibir tratamiento. J Coll Stud Psicoterapeuta. 2010; 24:49–58.

18. Beiter R, Nash R, McCrady M, Rhoades D, Linscomb M, et al. La prevalencia y los correlatos de la depresión, la ansiedad y el estrés en una muestra de estudiantes universitarios. J Afect Disord. 2015;173:90–6.

19. Nadeem M, Ali A, Buzdar MA. La asociación entre la religiosidad musulmana y la depresión, la ansiedad y el estrés de los estudiantes universitarios adultos jóvenes. J Relig Salud. 2017;56(4):1170–9.

20. Saeed H, Saleem Z, Ashraf M, et al. Determinantes de la ansiedad y la depresión entre estudiantes universitarios de Lahore. Int J Ment Heal Addict. 2018;16(5):1283–98.

21. Shamsuddin K, Fadzil F, Ismail WSW, Shah SA, Omar K, et al. Correlatos de depresión, ansiedad y estrés entre los estudiantes universitarios de Malasia. Psiquiatría asiática J. 2013, 6 (4): 318-23.

22. Taneja N, Sachdeva S, Dwivedi N. Evaluación de la depresión, la ansiedad y el estrés entre estudiantes de medicina matriculados en una facultad de medicina de Nueva Delhi, India. Indian J Soc Psiquiatría. 2018;34(2):157–62.

23. Brufaerts R, Mortier P, Kiekens G, Auerbach RP, et al. Problemas de salud mental en estudiantes universitarios de primer año: prevalencia y funcionamiento académico. J Afect Disord. 2018;225:97–103.

24. Alim SMA, Kibria HM, Islam SME, et al. Traducción de DASS 21 al bengali y validación entre estudiantes de medicina. Bangladesh J Psiquiatría. 2017;28(2):67–70.

25. Alim SMA, Rabbani HM, Karim MG, Mullick E, et al. Evaluación de la depresión, la ansiedad y el estrés entre los estudiantes de MBBS de primer año de una facultad de medicina pública, Bangladesh. Bangladesh J Psiquiatría. 2017;29(1):23–9.

26. Hossain MD, Ahmed HU, Chowdhury WA, Niessen LW, Alam DS. Trastornos mentales en Bangladesh: una revisión sistemática. BMC Psiquiatría. 2014;14(1):216.

27. Mamun MAA, Griffiths MD. La asociación entre adicto a Facebook ción y depresión: un estudio de encuesta piloto entre estudiantes de Bangladesh. Res. Psiquiatría. 2019;271:628–33.

28. Mamun MA, Raf MA, Hasan MZ, et al. Prevalencia y factores de riesgo psiquiátricos del uso excesivo de Internet entre estudiantes de posgrado que buscan trabajo en el norte de Bangladesh: un estudio piloto. Int J Ment Heal Addict. 2019. <https://doi.org/10.1007/s11469-019-00066-5>.

29. Mateen BA, Liley J, Denniston AK, Holmes CC, Vollmer SJ. Mejorar la calidad del aprendizaje automático en aplicaciones de salud e investigación clínica. Inteligencia de Nat Mach. 2020;2(10):554–6.

30. Roberts M, Driggs D, Thorpe M, Gilbey J, et al. Errores comunes y recomendaciones para usar el aprendizaje automático para detectar y pronosticar COVID-19 usando radiografías de tórax y tomografías computarizadas. Inteligencia de Nat Mach. 2021;3(3):199–217.

31. Topol EJ. Medicina de alto rendimiento: la convergencia de lo humano y lo inteligencia artificial. Nat Med. 2019, 25 (1): 44–56.

32. Organización Mundial de la Salud. Abordar los determinantes socioeconómicos de los hábitos alimentarios saludables y los niveles de actividad física entre los adolescentes. 2006, Venecia, Italia.

33. Chobanian AV, Bakris GL, Black HR. El séptimo informe de la reunión comité nacional de prevención, detección, evaluación y tratamiento de la presión arterial alta: el informe JNC 7. JAMA. 2003;289(19):2560–72.

34. Índice de enfermedades y condiciones- Hipotensión. Nacional del Corazón Pulmón y Instituto de la Sangre. 2008. http://www.nhlbi.nih.gov/health/dci/Diseases/hyp/hip_lo_que_es.html Consultado el 16 de septiembre de 2008.

35. Majed HT, Sadek AA. Prehipertensión e hipertensión en col. Estudiantes universitarios en Kuwait: un tema olvidado. J Fam Community Med. 2012;19(2):105.

36. Equipo central R. R: un lenguaje y entorno para la computación estadística. Viena: Fundación R para la Computación Estadística. <http://www.R-proyecto.org/>; 2013.

37. Kursu MB, Rudnicki WR. Selección de características con el paquete Boruta. Estado J suave 2010;36(11):1–13.

38. Igual L, Seguí S. Introducción a la Ciencia de Datos. Cham: Springer; 2017.

39. Nilsson NL. Introducción al aprendizaje automático, 1997, CA.

40. Breiman L. Bosques aleatorios. Aprender Mach. 2001;45(1):5–32.

- Springer Nature se mantiene neutral con respecto a los reclamos jurisdiccionales en mapas publicados y afiliaciones institucionales.

