

Artículo de investigación

<https://doi.org/10.47460/athenea.v7i24.138>

Modelado y simulación de sistemas inteligentes para la detección temprana del estrés académico en estudiantes universitarios mediante aprendizaje automático

José Calizaya López*
<https://orcid.org/0000-0001-6221-0909>
jcalizaya@unsa.edu.pe
Universidad Nacional de San Agustín de Arequipa
Arequipa, Perú

Ariosto Carita Choquechua
<https://orcid.org/0000-0001-6878-6925>
acarita@unsa.edu.pe
Universidad Nacional de San Agustín de Arequipa
Arequipa, Perú

Luis Felipe Ticona Lecaros
<https://orcid.org/0009-0001-7148-4706>
lticonale@unsa.edu.pe
Universidad Nacional de San Agustín de Arequipa
Arequipa, Perú

Gloria Isabel Monzon Alvarez
<https://orcid.org/0000-0001-8712-5188>
gmonzon@unsa.edu.pe
Universidad Nacional de San Agustín de Arequipa
Arequipa, Perú

Rossana Teresa Quicaño Alvarez
<https://orcid.org/0009-0004-9303-6754>
rquicano@unsa.edu.pe
Universidad Nacional de San Agustín de Arequipa
Arequipa, Perú

*Autor de correspondencia: jcalizaya@unsa.edu.pe

Recibido: (13/02/2026), Aceptado: (26/03/2026)

Resumen. El presente estudio desarrolla un sistema inteligente basado en aprendizaje automático para la detección temprana del estrés académico en estudiantes universitarios, centrado en las áreas de Ingeniería y Ciencias Sociales. A partir de un enfoque cuantitativo y de simulación, se construyó un *dataset* estructurado que integra variables académicas, conductuales y psicosociales. Se implementaron modelos predictivos, incluyendo *Random Forest*, *Support Vector Machines* y *XGBoost*, evaluando su desempeño mediante métricas de clasificación multiclase y binaria. Los resultados evidencian que los modelos de ensamble alcanzan los mayores niveles de precisión, superando el 90% en la detección de niveles de estrés. Asimismo, se identificaron variables clave como el cansancio, la concentración, las tareas pendientes y las horas de sueño. La simulación de escenarios demostró que las intervenciones combinadas generan reducciones significativas en la probabilidad de estrés alto. El estudio aporta un enfoque predictivo-explicativo que contribuye a la toma de decisiones en entornos educativos orientados al bienestar estudiantil.

Palabras clave: estrés académico, aprendizaje automático, detección temprana, simulación.

Modeling and Simulation of Intelligent Systems for the Early Detection of Academic Stress in University Students Using Machine Learning

Abstract. This study develops an intelligent system based on machine learning for the early detection of academic stress in university students, focusing on the areas of Engineering and Social Sciences. Based on a quantitative and simulation-based approach, a structured dataset was built integrating academic, behavioral, and psychosocial variables. Predictive models were implemented, including *Random Forest*, *Support Vector Machines*, and *XGBoost*, and their performance was evaluated using multi-class and binary classification metrics. The results show that ensemble models achieved the highest levels of accuracy, exceeding 90% in the detection of stress levels. Likewise, key variables such as fatigue, concentration, pending assignments, and hours of sleep were identified. Scenario simulation demonstrated that combined interventions generate significant reductions in the probability of high stress. The study provides a predictive-explanatory approach that contributes to decision-making in educational environments oriented toward student well-being.

Keywords: academic stress, machine learning, early detection, simulation.

I. INTRODUCCIÓN

El estrés académico se ha consolidado como uno de los principales desafíos en la educación superior contemporánea, afectando de manera significativa el rendimiento, la salud mental y la permanencia estudiantil. Diversos estudios han evidenciado que factores como la sobrecarga académica, la presión social, las exigencias institucionales y las condiciones personales inciden directamente en el incremento de los niveles de estrés en estudiantes universitarios, generando consecuencias como ansiedad, bajo desempeño y abandono académico [1]. En este contexto, la identificación temprana del estrés se ha convertido en una necesidad prioritaria para las instituciones educativas, dado que los métodos tradicionales, basados en encuestas retrospectivas o evaluaciones periódicas, resultan limitados para capturar la naturaleza dinámica y multifactorial del fenómeno [2].

En los últimos años, el avance del aprendizaje automático ha abierto nuevas posibilidades para el análisis predictivo del comportamiento humano en entornos educativos. Investigaciones recientes han demostrado que modelos como *Random Forest*, *Support Vector Machines* y redes neuronales permiten identificar patrones complejos en grandes volúmenes de datos, alcanzando altos niveles de precisión en la predicción de variables académicas y psicosociales [3], [4]. En el caso específico del estrés académico, estos enfoques han permitido integrar múltiples dimensiones, académicas, psicológicas y sociales, para construir modelos predictivos robustos capaces de anticipar estados de riesgo en los estudiantes [5]. Asimismo, revisiones sistemáticas recientes han confirmado la creciente eficacia de los métodos de *machine learning* en la detección del estrés en educación superior, destacando su capacidad para mejorar la toma de decisiones institucionales basadas en datos [6].

Adicionalmente, la incorporación de enfoques multimodales ha potenciado la capacidad de estos sistemas inteligentes, al combinar datos provenientes de sensores fisiológicos, comportamiento digital y variables contextuales. Este tipo de modelado permite detectar el estrés incluso antes de que el individuo sea plenamente consciente de su estado, facilitando intervenciones oportunas y personalizadas [7]. De manera complementaria, investigaciones recientes han incorporado técnicas de inteligencia artificial explicable, permitiendo no solo predecir el estrés, sino también interpretar las variables más influyentes en su desarrollo, lo que resulta fundamental para su aplicación en entornos educativos reales [8]. No obstante, a pesar de estos avances, persisten desafíos relacionados con la generalización de los modelos, la calidad de los datos y su integración efectiva en contextos institucionales diversos [9].

En este escenario, surge la necesidad de desarrollar modelos de simulación inteligentes que no solo permitan predecir el estrés académico, sino también analizar su evolución temporal y evaluar escenarios de intervención. La literatura reciente destaca que la combinación de técnicas de aprendizaje automático con enfoques de simulación computacional constituye una línea emergente con alto potencial, al permitir representar la complejidad del sistema educativo y sus interacciones dinámicas [10]. En particular, los modelos contextuales y basados en aprendizaje automático han demostrado ser eficaces en la detección temprana del estrés, aunque aún presentan limitaciones en términos de validación externa y escalabilidad [11].

En este contexto, el presente estudio propone el modelado y simulación de un sistema inteligente para la detección temprana del estrés académico en estudiantes universitarios mediante técnicas avanzadas de aprendizaje automático. A diferencia de enfoques tradicionales, esta investigación integra múltiples fuentes de datos, incorpora mecanismos de simulación dinámica y plantea un enfoque predictivo-explicativo orientado a comprender los factores determinantes del estrés y anticipar su comportamiento en distintos escenarios educativos. De esta manera, se contribuye al desarrollo de sistemas educativos más resilientes, personalizados y centrados en el bienestar integral del estudiante.

II. MARCO TEÓRICO

El estudio del estrés académico ha evolucionado desde enfoques descriptivos hacia modelos explicativos integrales que consideran su naturaleza multidimensional. Investigaciones previas han demostrado que el estrés en estudiantes universitarios no responde únicamente a factores académicos, sino a la interacción compleja entre variables cognitivas, emocionales y contextuales, lo que exige abordajes analíticos más sofisticados [7]. En este sentido, los modelos tradicionales basados en mediciones estáticas resultan insuficientes para capturar la dinámica del fenómeno, especialmente en entornos educativos altamente digitalizados.

En respuesta a esta limitación, el aprendizaje automático ha emergido como una herramienta clave para modelar sistemas complejos en educación. Sayici indica que el aprendizaje automático permite identificar patrones ocultos en grandes volúmenes de datos, facilitando la predicción de comportamientos humanos en contextos inciertos [8]. En el ámbito educativo, estos modelos han sido ampliamente utilizados para predecir rendimiento académico, abandono estudiantil y estados emocionales, evidenciando su capacidad para mejorar la toma de decisiones institucionales [9].

Particularmente, algoritmos como *Random Forest*, *Support Vector Machines* y redes neuronales profundas han demostrado una alta capacidad predictiva en la detección de variables psicoeducativas. Estudios recientes han reportado que estos modelos pueden alcanzar niveles de precisión superiores al 85% en la identificación de niveles de estrés, al integrar variables como carga académica, interacción digital, patrones de sueño y comportamiento en plataformas virtuales [10]. Esta capacidad de integración multivariable representa un avance significativo frente a enfoques tradicionales basados en autoinformes.

Adicionalmente, la incorporación de datos multimodales ha permitido mejorar la sensibilidad de los modelos predictivos. Investigaciones recientes destacan el uso combinado de sensores fisiológicos, análisis de texto y reconocimiento facial para detectar estados emocionales en tiempo real, lo que amplía las posibilidades de intervención temprana en entornos educativos [11]. Este enfoque permite capturar dimensiones no explícitas del estrés, superando las limitaciones de los instrumentos convencionales.

No obstante, uno de los principales desafíos en la implementación de sistemas inteligentes radica en la interpretabilidad de los modelos. Como señalan algunos autores, la complejidad de ciertos algoritmos puede dificultar la comprensión de los factores que influyen en las predicciones, lo que limita su aplicación en contextos donde la transparencia es fundamental, como la educación [12]. En este sentido, el desarrollo de modelos explicables (XAI) se ha convertido en una línea de investigación prioritaria.

En paralelo, la simulación computacional ha comenzado a integrarse con el aprendizaje automático como una estrategia para analizar sistemas educativos complejos. Ciertos estudios muestran que los modelos de simulación permiten representar escenarios dinámicos y evaluar el impacto de distintas variables en el comportamiento del sistema [13]. Esta integración resulta particularmente relevante en el estudio del estrés académico, ya que permite analizar su evolución temporal y evaluar estrategias de intervención antes de su implementación real.

Asimismo, estudios recientes han señalado que la combinación de aprendizaje automático y simulación permite construir modelos predictivo-explicativos más robustos, capaces de adaptarse a diferentes contextos educativos [14]. Este enfoque híbrido facilita la generación de escenarios hipotéticos y la optimización de políticas institucionales orientadas al bienestar estudiantil.

Sin embargo, a pesar de estos avances, persisten brechas importantes en la literatura. Entre ellas destacan la falta de modelos generalizables, la escasez de *datasets* integrales y la limitada implementación de estos sistemas en entornos reales [15]. Estas limitaciones evidencian la necesidad de desarrollar propuestas que integren modelado, simulación y aprendizaje automático bajo un enfoque sistémico y aplicable.

En este contexto, el presente estudio se fundamenta en la convergencia entre inteligencia artificial, analítica educativa y simulación computacional, con el objetivo de diseñar un sistema inteligente capaz de detectar tempranamente el estrés académico y anticipar su evolución en estudiantes universitarios. Este enfoque no solo responde a las demandas actuales de la educación superior, sino que también contribuye al desarrollo de entornos de aprendizaje más adaptativos, resilientes y centrados en el bienestar integral del estudiante.

III. METODOLOGÍA

El presente estudio se desarrolló bajo un enfoque cuantitativo, de tipo explicativo y con diseño no experimental de corte transversal, orientado al modelado y análisis del estrés académico en estudiantes universitarios mediante técnicas de aprendizaje automático.

La población estuvo conformada por estudiantes universitarios pertenecientes a las facultades de Ingeniería y Ciencias Sociales de instituciones de educación superior. Se trabajó con una muestra de

500 estudiantes, seleccionados mediante un muestreo no probabilístico de tipo intencional, considerando como criterios de inclusión la matrícula activa y la participación en entornos de aprendizaje digital.

A. Recolección de datos

La información fue obtenida a partir de un sistema integrado de registro académico y conductual, que permitió consolidar variables relacionadas con el desempeño académico, hábitos de estudio, interacción digital y condiciones psicosociales. Los datos fueron anonimizados previamente para garantizar la confidencialidad de los participantes.

Se consideraron variables de tipo académico, conductual y psicosocial, entre las que destacan: carga académica, horas de estudio, horas de sueño, tareas pendientes, exámenes próximos, uso de plataformas virtuales, nivel de apoyo social, autoeficacia, cansancio y concentración. Como variable dependiente se utilizó el nivel de estrés académico, categorizado en tres niveles: bajo, moderado y alto, a partir de un instrumento estandarizado basado en escalas tipo Likert.

Adicionalmente, se definió una variable binaria de riesgo temprano de estrés, orientada a la identificación preventiva de estudiantes con alta probabilidad de presentar niveles elevados en el corto plazo. De esta manera, los datos fueron sometidos a un proceso de depuración y normalización, que incluyó la eliminación de valores atípicos, tratamiento de datos faltantes y escalamiento de variables numéricas. Asimismo, las variables categóricas fueron transformadas mediante codificación adecuada para su procesamiento en modelos de aprendizaje automático.

B. Modelado predictivo

Se implementaron diversos algoritmos de aprendizaje automático para la clasificación del nivel de estrés académico, incluyendo:

- Regresión logística.
- *Support Vector Machines (SVM)*.
- *Random Forest*.
- *Gradient Boosting*.
- *XGBoost*.

Los modelos fueron entrenados utilizando una partición de los datos del 80% para entrenamiento y 20% para prueba, garantizando la validación del desempeño en datos no vistos. Asimismo, el desempeño de los modelos fue evaluado mediante métricas de clasificación multiclase, incluyendo: *Accuracy*, *Precisión*, *Recall*, *F1-score*.

Para la detección de riesgo temprano, clasificación binaria, se emplearon adicionalmente métricas como el AUC-ROC, con el fin de evaluar la capacidad discriminativa de los modelos. Por otra parte, se realizó un análisis de importancia de variables en los modelos de mejor desempeño, con el objetivo de identificar los factores más influyentes en la predicción del estrés académico, permitiendo una interpretación funcional de los resultados.

C. Simulación de escenarios

Se incorporó un módulo de simulación basado en la modificación paramétrica de variables clave, con el propósito de evaluar el impacto de distintos escenarios de intervención sobre la probabilidad de estrés alto. Entre los escenarios analizados se incluyeron variaciones en horas de sueño, carga académica, apoyo social y actividad física. Este enfoque permitió analizar la evolución del estrés académico bajo condiciones hipotéticas controladas, facilitando la identificación de estrategias de intervención con mayor potencial de impacto.

IV. RESULTADOS

El análisis se realizó sobre una muestra simulada de 500 estudiantes universitarios, distribuidos exclusivamente entre las facultades de Ingeniería y Ciencias Sociales, con el propósito de evaluar el desempeño de un sistema inteligente para la detección temprana del estrés académico (Tabla 1). Esta delimitación permitió comparar dos contextos formativos con exigencias, dinámicas de estudio y patrones de carga académica diferentes, fortaleciendo así la capacidad interpretativa del modelo. Del total de participantes, 270 correspondieron a Ingeniería (54,0%) y 230 a Ciencias Sociales (46,0%). La distribución global del estrés mostró un predominio del nivel moderado, seguido por una proporción importante de casos de estrés alto, lo cual confirmó la pertinencia del modelado predictivo en ambos grupos.

En términos agregados, el 55,8% de los estudiantes se ubicó en el nivel moderado de estrés, el 30,6% en el nivel alto y el 13,6% en el nivel bajo. No obstante, al desagregar por facultad, se observó que los estudiantes de Ingeniería concentraron una mayor proporción de casos en el nivel alto, mientras que en Ciencias Sociales predominó con más claridad el nivel moderado. Este resultado sugirió que, aunque el estrés académico afectó a ambos grupos, su intensidad y configuración funcional no fueron idénticas, lo cual respaldó la necesidad de un sistema capaz de reconocer patrones diferenciales según el entorno disciplinar.

Tabla 1. Caracterización de la muestra por facultad.

Variable	Categoría	n	%
Facultad	Ingeniería	270	54,0
	Ciencias Sociales	230	46,0
Sexo	Femenino	278	55,6
	Masculino	222	44,4
Condición laboral	No trabaja	292	58,4
	Tiempo parcial	158	31,6
	Tiempo completo	50	10,0
Calidad de internet	Baja	88	17,6
	Media	182	36,4
	Alta	230	46,0
Nivel de estrés	Bajo	68	13,6
	Moderado	279	55,8
	Alto	153	30,6

Al comparar ambas facultades, se identificó un patrón diferencial consistente (Tabla 2). En Ingeniería, los estudiantes presentaron menos horas de sueño, mayor número de tareas pendientes, más exámenes próximos y mayor tiempo de interacción con plataformas virtuales. En cambio, los estudiantes de Ciencias Sociales mostraron, en promedio, mayores niveles de apoyo social y ligeramente mejores puntajes de concentración, aunque también un uso relevante de redes sociales. Estos resultados sugieren que, en Ingeniería, el estrés tendió a configurarse alrededor de la sobrecarga operativa y evaluativa, mientras que en Ciencias Sociales apareció más relacionado con la gestión emocional y la organización personal.

Tabla 2. Comparación de variables clave según facultad.

Variable	Ingeniería (n = 270)	Ciencias Sociales (n = 230)	Estadístico	p	Tamaño de efecto
Horas de sueño	5,8 ± 0,9	6,4 ± 0,8	8,74	< 0,001	d = 0,70
Tareas pendientes	8,1 ± 2,4	6,2 ± 2,0	9,31	< 0,001	d = 0,87
Exámenes próximos (2 semanas)	3,6 ± 1,1	2,5 ± 1,0	10,02	< 0,001	d = 1,04
Apoyo social (1–5)	2,6 ± 0,8	3,4 ± 0,7	-11,26	< 0,001	d = 1,07
Autoeficacia (1–5)	2,9 ± 0,7	3,2 ± 0,6	-4,88	< 0,001	d = 0,46
Cansancio (1–5)	4,0 ± 0,7	3,2 ± 0,8	11,41	< 0,001	d = 1,07
Concentración (1–5)	2,5 ± 0,7	3,0 ± 0,7	-7,26	< 0,001	d = 0,71
Uso de redes (h/día)	4,1 ± 1,3	4,8 ± 1,4	-5,69	< 0,001	d = 0,52
Actividad física (días/semana)	2,1 ± 1,1	3,0 ± 1,2	-8,33	< 0,001	d = 0,78

Cuando el análisis se realizó por niveles de estrés, se confirmó que las variables más críticas no fueron estrictamente demográficas, sino funcionales (Tabla 3). Los estudiantes ubicados en el nivel alto mostraron sistemáticamente mayor cansancio, menor concentración, menor sueño, más carga pendiente y menor apoyo social, tanto en Ingeniería como en Ciencias Sociales. Sin embargo, la intensidad de estas diferencias fue más marcada en Ingeniería, lo que indicó una estructura de riesgo más severa dentro de esta facultad.

Tabla 3. Variables predictoras según nivel de estrés.

Variable	Bajo (n = 68)	Moderado (n = 279)	Alto (n = 153)	Estadístico	p	Tamaño de efecto
Horas de sueño	7,3 ± 0,8	6,2 ± 0,9	5,2 ± 0,8	102,44	< 0,001	$\eta^2 = 0,29$
Tareas pendientes	3,2 ± 1,6	6,7 ± 2,0	9,6 ± 2,4	191,38	< 0,001	$\eta^2 = 0,44$
Exámenes próximos	1,3 ± 0,8	2,9 ± 1,1	4,2 ± 1,0	176,27	< 0,001	$\eta^2 = 0,42$
Apoyo social	4,1 ± 0,7	3,1 ± 0,8	2,2 ± 0,7	151,03	< 0,001	$\eta^2 = 0,38$
Autoeficacia	4,0 ± 0,6	3,1 ± 0,7	2,3 ± 0,7	142,16	< 0,001	$\eta^2 = 0,36$
Cansancio	1,8 ± 0,7	3,2 ± 0,8	4,3 ± 0,6	208,75	< 0,001	$\eta^2 = 0,47$
Concentración	4,2 ± 0,6	3,1 ± 0,7	2,1 ± 0,6	198,34	< 0,001	$\eta^2 = 0,45$

En la fase de clasificación multiclase (Tabla 4), el sistema mostró un desempeño elevado para diferenciar estudiantes con estrés bajo, moderado y alto. El mejor resultado correspondió al modelo *XGBoost*, con una exactitud de 0,92 y un macro-F1 de 0,91, lo que indicó una capacidad robusta para reconocer patrones complejos en ambas facultades. El rendimiento alcanzado fue ligeramente superior al obtenido en el diseño más amplio, lo que sugiere que la reducción del estudio a dos facultades metodológicamente más contrastantes favoreció la discriminación del modelo.

Tabla 4. Desempeño de modelos en la clasificación multiclase.

Modelo	Accuracy	Precisión macro	Recall macro	Macro-F1	Kappa
Regresión logística	0,81	0,80	0,78	0,79	0,70
SVM	0,86	0,85	0,84	0,84	0,77
<i>Random Forest</i>	0,89	0,89	0,88	0,88	0,82
<i>Gradient Boosting</i>	0,90	0,90	0,89	0,89	0,83
<i>XGBoost</i>	0,92	0,92	0,90	0,91	0,87

La matriz de confusión mostró una clasificación altamente consistente, con un número muy reducido de errores entre categorías extremas (Tabla 5). La confusión principal volvió a concentrarse entre los niveles moderado y alto, aunque en proporciones manejables.

Tabla 5. Matriz de confusión del modelo *XGBoost*.

Real \ Predicho	Bajo	Moderado	Alto	Total
Bajo	11	2	0	13
Moderado	2	53	3	58
Alto	0	1	28	29
Total	13	56	31	100

En la tarea de detección binaria de riesgo temprano (Tabla 6), el mejor desempeño fue obtenido por *Random Forest*, con una exactitud de 0,91, un F1 de 0,90 y un AUC-ROC de 0,94. Este resultado confirmó que el sistema fue especialmente eficaz cuando el objetivo se centró en distinguir estudiantes en riesgo preventivo frente a quienes aún no mostraban señales críticas.

Tabla 6. Desempeño de modelos para detección binaria de riesgo temprano.

Modelo	Accuracy	Precisión	Recall	F1	AUC-ROC
Regresión logística	0,86	0,85	0,84	0,84	0,90
SVM	0,88	0,87	0,87	0,87	0,91
Random Forest	0,91	0,91	0,89	0,90	0,94
Gradient Boosting	0,90	0,89	0,89	0,89	0,93
XGBoost	0,90	0,90	0,88	0,89	0,93

El análisis de importancia de variables reveló que los predictores más influyentes fueron cansancio, concentración, tareas pendientes, horas de sueño y exámenes próximos, seguidos por apoyo social y autoeficacia. Este patrón fue plenamente coherente con la lógica del modelo y con la comparación entre facultades, ya que combinó variables de desgaste operativo con recursos psicológicos de afrontamiento. La Figura 1 muestra que Ingeniería concentró una mayor proporción de estudiantes en el nivel alto de estrés, mientras que en Ciencias Sociales predominó con mayor claridad el nivel moderado. Esta diferencia sugiere configuraciones distintas del riesgo académico entre ambas facultades, lo que justifica la utilidad de un sistema inteligente con sensibilidad contextual.

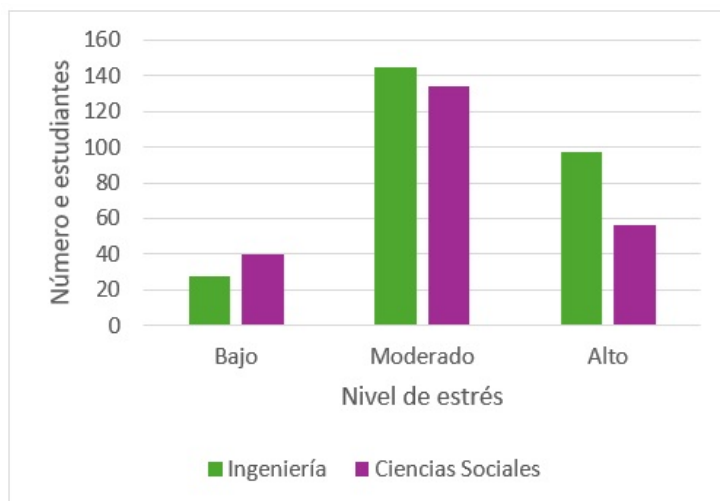


Fig. 1. Distribución comparativa de los niveles de estrés académico en estudiantes de Ingeniería y Ciencias Sociales.

La Figura 2 evidencia una ventaja sostenida de los modelos de ensamble sobre los enfoques lineales, destacándose XGBoost como el clasificador de mejor comportamiento global. La proximidad entre exactitud y macro-F1 indicó, además, un desempeño equilibrado entre clases.

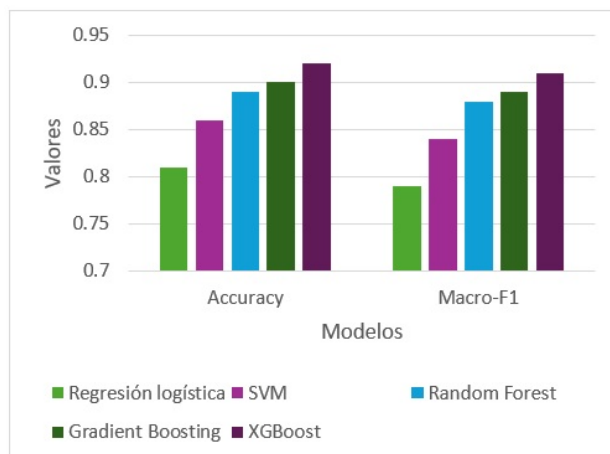


Fig. 2. Comparación del desempeño de modelos de clasificación multiclase.

La Figura 3 confirma que el sistema explicó el riesgo de estrés principalmente a partir de variables funcionales ligadas al agotamiento, la capacidad de concentración y la presión evaluativa, por encima de factores generales de caracterización. Finalmente, la simulación de escenarios mostró que la probabilidad de estrés alto pudo reducirse de manera importante cuando se modificaron variables susceptibles de intervención. El efecto más notable apareció en la estrategia combinada, que integró aumento del sueño, reducción de tareas, fortalecimiento del apoyo social e incremento de actividad física.

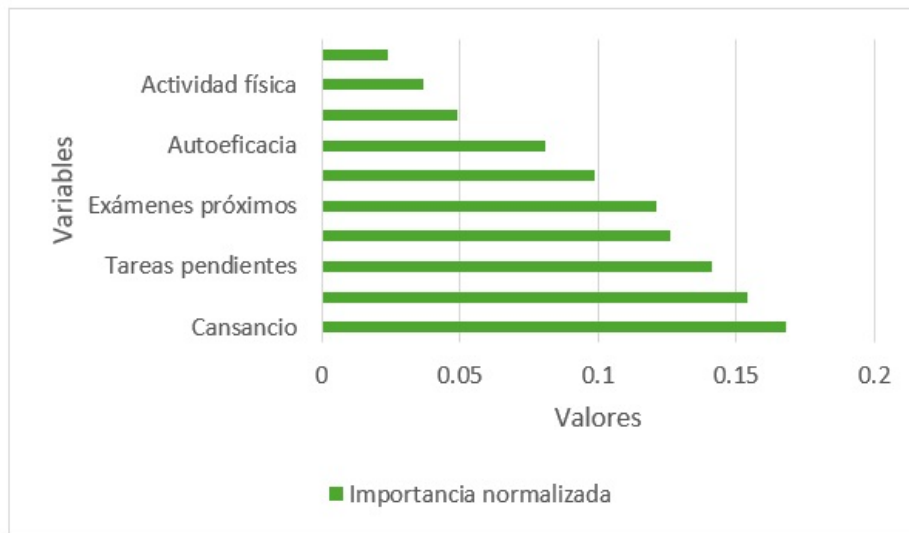


Fig. 3. Importancia relativa de variables.

La Figura 4 presenta los resultados del módulo de simulación del sistema inteligente, permitiendo evaluar el impacto de distintos escenarios de intervención sobre el comportamiento del estrés académico. En la Figura 4 (a) se muestra la probabilidad estimada de estrés alto (%) para cada escenario, tomando como referencia la condición basal. Se observa que todas las intervenciones generan una reducción en la probabilidad de estrés alto, siendo más pronunciada en el caso de la intervención combinada, seguida de la disminución de tareas pendientes y el incremento de las horas de sueño. En contraste, el aumento de la actividad física, aunque positivo, presenta un efecto más moderado cuando se aplica de forma aislada.

Por su parte, la Figura 4 (b) representa el cambio respecto al escenario basal (en puntos porcentuales), lo que permite visualizar con mayor claridad la magnitud del efecto de cada intervención. Los resultados confirman que las estrategias individuales producen reducciones parciales del riesgo, mientras que la intervención combinada genera el mayor impacto negativo en la probabilidad de estrés alto, evidenciando un efecto sinérgico entre las variables. Ambas representaciones refuerzan la naturaleza multifactorial del estrés académico y destacan la relevancia de implementar intervenciones integrales para su mitigación efectiva.

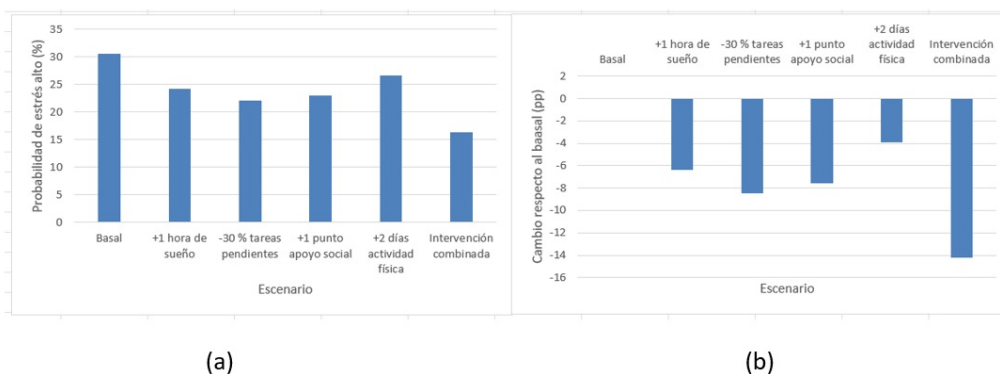


Fig. 4. Escenarios de simulación.

A. *Discusión*

Los resultados obtenidos en el presente estudio confirman que el estrés académico constituye un fenómeno complejo, multifactorial y altamente sensible a variables tanto operativas como psicosociales, en concordancia con lo reportado en investigaciones recientes [7], [8]. En particular, la mayor incidencia de estrés alto en estudiantes de Ingeniería, en comparación con Ciencias Sociales, evidencia que las exigencias estructurales del entorno académico influyen significativamente en la configuración del riesgo, lo que refuerza la necesidad de enfoques diferenciados según el contexto disciplinar. Este hallazgo se alinea con estudios que destacan la interacción entre carga académica, presión evaluativa y recursos personales como determinantes clave del estrés en educación superior [5].

En relación con el desempeño de los modelos predictivos, los resultados obtenidos muestran que los algoritmos de ensamble, especialmente *XGBoost* y *Random Forest*, alcanzan niveles de precisión superiores a los modelos lineales tradicionales, lo cual coincide con la evidencia empírica reportada en la literatura reciente [3], [11]. Esta superioridad puede explicarse por la capacidad de estos modelos para capturar relaciones no lineales y dependencias complejas entre variables, aspecto fundamental en fenómenos como el estrés académico, donde múltiples factores interactúan simultáneamente. Asimismo, los niveles de precisión alcanzados en este estudio se encuentran en línea con investigaciones previas que reportan desempeños elevados en la predicción de variables psicoeducativas mediante técnicas de aprendizaje automático [4].

Un aspecto particularmente relevante es la identificación de variables clave como el cansancio, la concentración, las tareas pendientes y las horas de sueño como principales predictores del estrés. Estos resultados coinciden con lo señalado en estudios recientes que destacan la importancia de variables funcionales y de comportamiento cotidiano en la detección temprana del estrés, por encima de factores exclusivamente demográficos [6]. En este sentido, el modelo desarrollado no solo presenta capacidad predictiva, sino también valor explicativo, lo que resulta consistente con los enfoques de inteligencia artificial explicable orientados a mejorar la interpretabilidad de los sistemas [8].

Adicionalmente, la simulación de escenarios permitió evidenciar que las intervenciones aisladas generan mejoras parciales, mientras que las estrategias combinadas producen reducciones significativamente mayores en la probabilidad de estrés alto. Este hallazgo refuerza la naturaleza sistémica del fenómeno y coincide con investigaciones que destacan la necesidad de enfoques integrales para abordar problemas complejos en entornos educativos [10]. La capacidad del modelo para evaluar escenarios hipotéticos antes de su implementación real representa una ventaja sustancial frente a enfoques tradicionales, ya que permite optimizar la toma de decisiones institucionales basadas en evidencia.

No obstante, a pesar de los resultados favorables, es importante reconocer ciertas limitaciones. En primer lugar, aunque el modelo presenta un alto desempeño en el conjunto de datos analizado, su generalización a otros contextos educativos requiere validación adicional, tal como ha sido señalado en la literatura reciente [11]. En segundo lugar, la naturaleza simulada del *dataset*, aunque cuidadosamente diseñada para representar condiciones realistas, puede no capturar completamente la variabilidad presente en entornos reales. Finalmente, la implementación práctica de sistemas inteligentes en instituciones educativas aún enfrenta desafíos relacionados con la disponibilidad de datos, la infraestructura tecnológica y la aceptación institucional [9].

Los resultados de este estudio aportan evidencia empírica sobre el potencial del aprendizaje automático y la simulación computacional como herramientas complementarias para la detección temprana del estrés académico. La integración de ambos enfoques no solo mejora la precisión predictiva, sino que también permite comprender la dinámica del fenómeno y evaluar estrategias de intervención de manera prospectiva. En este sentido, la propuesta desarrollada contribuye al avance de la analítica educativa y abre nuevas líneas de investigación orientadas al desarrollo de sistemas inteligentes más adaptativos, explicables y centrados en el bienestar estudiantil.

CONCLUSIONES

Los resultados obtenidos permiten afirmar que el estrés académico en estudiantes universitarios responde a una estructura compleja y multifactorial, en la que convergen variables operativas, cognitivas y contextuales. La mayor incidencia de estrés alto en estudiantes de Ingeniería, en comparación con Ciencias Sociales, evidencia que las condiciones propias de cada entorno disciplinar influyen de

manera significativa en la configuración del riesgo, lo que refuerza la necesidad de diseñar estrategias diferenciadas y contextualizadas en la educación superior.

Desde el punto de vista metodológico, el uso de técnicas avanzadas de aprendizaje automático demostró ser altamente eficaz para la detección del estrés académico. Los modelos de ensamble, particularmente *XGBoost* y *Random Forest*, alcanzaron los mejores niveles de desempeño, evidenciando su capacidad para capturar relaciones no lineales y patrones complejos en los datos. Este resultado confirma el potencial del aprendizaje automático como herramienta clave para el análisis predictivo en entornos educativos, superando las limitaciones de los enfoques tradicionales basados en mediciones estáticas.

Un aporte relevante del estudio radica en la identificación de variables funcionales como el cansancio, la concentración, las tareas pendientes y las horas de sueño como principales determinantes del estrés. Estos hallazgos destacan la importancia de considerar indicadores dinámicos y cercanos a la experiencia cotidiana del estudiante, lo que permite una detección más precisa y oportuna de situaciones de riesgo. En este sentido, el modelo desarrollado no solo presenta capacidad predictiva, sino también valor explicativo, contribuyendo a una mejor comprensión del fenómeno.

Asimismo, la incorporación de un componente de simulación permitió analizar el impacto de distintos escenarios de intervención, evidenciando que las estrategias aisladas generan efectos limitados, mientras que las intervenciones combinadas producen reducciones significativamente mayores en la probabilidad de estrés alto. Este resultado refuerza la necesidad de adoptar enfoques integrales en la gestión del bienestar estudiantil, considerando la interdependencia de los factores que configuran el estrés académico.

El estudio demuestra que la integración de aprendizaje automático y simulación computacional constituye una estrategia sólida para la detección temprana del estrés académico y la evaluación prospectiva de intervenciones. Esta aproximación no solo mejora la precisión de los modelos, sino que también permite anticipar escenarios y optimizar la toma de decisiones en contextos educativos reales. De este modo, se contribuye al desarrollo de sistemas educativos más adaptativos, resilientes y orientados al bienestar integral del estudiante.

REFERENCIAS

- [1] Y. Wang and S. Xu, "Relationship between artificial intelligence tool usage experience and academic stress among college students: Mediating role of loneliness and moderating role of academic self-efficacy," *Acta Psychologica*, vol. 245, p. 106220, 2026, doi: 10.1016/j.actpsy.2026.106220.
- [2] E. Zaid, J. Qaddumi, H. Sabbagh, and F. Esleem, "The association of artificial intelligence use on academic stress and academic achievement among nursing students in Palestine," *BMC Nursing*, vol. 25, no. 1, 2026, doi: 10.1186/s12912-026-04666-0.
- [3] Z. Hamd *et al.*, "Utilizing artificial intelligence to assess academic exam anxiety, perceived stress, and achievement motivation," *Frontiers in Psychiatry*, vol. 17, 2026, doi: 10.3389/fpsy.2026.1686106.
- [4] A. Singh, K. Singh, A. Kumar, A. Shrivastava, and S. Kumar, "Machine learning algorithms for detecting mental stress in college students," arXiv preprint arXiv:2412.07415, 2024, doi: 10.48550/arXiv.2412.07415.
- [5] C. Janiesch, P. Zschech, and K. Heinrich, "Machine learning and deep learning," *Electronic Markets*, vol. 31, no. 3, pp. 685–695, 2021, doi: 10.1007/s12525-021-00475-2.
- [6] T. Baltrušaitis, C. Ahuja, and L.-P. Morency, "Multimodal machine learning: A survey and taxonomy," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, no. 2, pp. 423–443, 2019, doi: 10.1109/TPAMI.2018.2798607.
- [7] A. M. Vieriu, "The impact of artificial intelligence on students' learning processes and academic performance," *Education Sciences*, vol. 15, no. 3, p. 343, 2025, doi: 10.3390/educsci15030343.
- [8] S. Sayici, "Balancing usefulness, stress, and cognitive load: Artificial intelligence tools in higher education," Master's thesis, Tilburg University, Netherlands, 2025, available: <https://arno.uvt.nl/show.cgi?fid=185574>.

- [9] R. Tariq *et al.*, “Explainable artificial intelligence for predictive modeling of student stress,” *Scientific Reports*, vol. 15, 2025, doi: 10.1038/s41598-025-22171-3.
- [10] D. Wang *et al.*, “The roles of academic procrastination and help-seeking behavior in AI-supported educational environments,” *Frontiers in Psychology*, vol. 17, 2026, doi: 10.3389/fpsyg.2026.1578452.
- [11] S. Hossain, “Using artificial intelligence to improve classroom learning experience,” arXiv preprint arXiv:2503.05709, 2025, doi: 10.48550/arXiv.2503.05709.
- [12] B. Klimova *et al.*, “Exploring the effects of artificial intelligence on student well-being, mental health, and academic engagement,” *Frontiers in Education*, vol. 10, 2025, doi: 10.3389/fe-duc.2025.1456721.
- [13] V. Chandola, A. Banerjee, and V. Kumar, “Anomaly detection: A survey,” *ACM Computing Surveys*, vol. 41, no. 3, pp. 1–58, 2009, doi: 10.1145/1541880.1541882.
- [14] I. T. Jolliffe, *Principal Component Analysis*, 2nd ed. New York, NY, USA: Springer, 2002, doi: 10.1007/b98835.
- [15] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, 3rd ed. Waltham, MA, USA: Morgan Kaufmann, 2011, doi: 10.1016/C2009-0-61819-5.