

Análisis de la Influencia de Factores Demográficos y Sociales en la Salud Mental de Diferentes Grupos Poblacionales



UNIVERSITAT
ROVIRA I VIRGILI



Universitat
Oberta
de Catalunya

Juan Andrés Girón Caballero

**MU Ingeniería Computacional y Matemática
Área de Inteligencia Artificial**

Dr. Antonio Sarasa

Nombre Tutor/a de TF

Profesor/a responsable de la asignatura

Fecha Entrega

9 de Junio de 2025

**Firma del director autorizando la entrega final del
TFM**

Título del trabajo:	Análisis de la Influencia de Factores Demográficos y Sociales en la Salud Mental de Diferentes Grupos Poblacionales
Nombre del autor:	Ing. Juan Andrés Girón Caballero
Nombre del consultor/a:	Dr. Antonio Sarasa
Nombre del PRA:	
Fecha de entrega (mm/aaaa):	Junio/2025
Titulación o programa:	MU Ingeniería Computacional y Matemática
Área del Trabajo Final:	Inteligencia Artificial
Idioma del trabajo:	Castellano
Palabras clave	Salud mental, XAI, factores sociales
Resumen del Trabajo	
<p>En este trabajo se analizan datos sobre salud mental incorporando factores sociales y ambientales, con el fin de desarrollar un modelo de clasificación que prediga la probabilidad de padecimiento. Se preprocesan los datos, se aplica reducción de dimensiones y se incorporan modelos de clusterización y clasificación que facilitan la identificación de patrones relevantes. Se emplea una metodología en cascada para organizar el proceso en análisis, diseño, implementación, verificación y mantenimiento, lo que permite ajustar el modelo según las necesidades del contexto. Integrar múltiples conjuntos de datos aumenta la robustez del sistema y ayuda a ofrecer una perspectiva más completa. El resultado es una herramienta útil y accesible para mejorar la toma de decisiones clínicas en el ámbito de la salud mental.</p>	

1 Tabla de contenido

2	Introducción	5
2.1.1	Contexto y justificación del trabajo	5
2.2	Objetivos del trabajo	6
2.3	Objetivos específicos	6
2.4	Impacto en sostenibilidad, ético-social y de diversidad	7
2.4.1	ODS 9: Industria, Innovación e Infraestructura	7
2.4.2	ODS 16: Paz, Justicia e Instituciones Sólidas	8
2.4.3	ODS 8: Trabajo Decente y Crecimiento Económico y ODS 10: Reducción de las Desigualdades	8
2.4.4	Enfoque y método seguido	8
2.4.5	Fase de Análisis:	8
2.4.6	Entregables e hitos	9
3	Estado del arte	12
4	Materiales y métodos	14
4.1	Fuentes de datos y criterios de selección	14
5	Clasificación metodológica de las variables	15
5.1	Clasificación conceptual de las variables	15
5.2	Justificación de la selección de variables	16
5.2.1	Variables consideradas	17
5.3	Procesamiento y análisis de datos	17
5.4	Desarrollo de modelos predictivos	20
5.4.1	Diferencias entre Regresión Logística, Random Forest y XGBoost	23
5.5	Evaluación e interpretabilidad de modelos	23
5.6	Implementación técnica y repositorio	24
5.6.1	Reflexión sobre el proceso	25
6	Resultados y discusión	26
6.1	Resultados iniciales	26
6.2	Justificación de la ampliación de variables y feature engineering	31
6.3	Comparativa Global y Síntesis	36
6.4	Resultados con explicabilidad	38

6.5	Resultados sobre conjunto alternativo de datos.....	41
7	Conclusiones y futuros trabajos	42
8	Anexos.....	44
9	Bibliografía.....	47
	Figura 1: Histograma de grupos poblacionales	19
	Figura 2: comparacion por region vs ansiedad	20
	Figura 3, ejemplo de regresion logistica	21
	Figura 4: ejemplo de arboles de decisiones	22
	Figura 5: ejemplo de XGB boost	22
	Figura 6: resultados de random forest basico	26
	Figura 7: histograma de clases.....	27
	Figura 8: random forest con ajuste de threshold	28
	Figura 9: resultado de regresion logistica con threshold	29
	Figura 10: XGboost tuned con balanceo de clases.....	30
	Figura 11: random forest básico con nuevos features	32
	Figura 12: random forest con threshold	32
	Figura 13: rregresion logistica con nuevos features.....	33
	Figura 14: XGBoost con threshold	34
	Figura 15: Ramdom forest con oversampling.....	34
	Figura 16: XGboost forest con oversampling	35
	Figura 17: XGboost forest tuneada con oversampling.....	35
	Figura 18: curva ROC XGBoost tunned.....	36

2 Introducció

2.1.1 Contexto y justificación del trabajo

La salud mental constituye uno de los pilares fundamentales para el bienestar integral de las personas y las sociedades. No obstante, la magnitud real de los trastornos mentales a nivel global, especialmente la ansiedad y la depresión, continúa siendo subestimada, en gran parte debido al estigma social, la falta de acceso a servicios especializados y la ausencia de sistemas eficientes de detección temprana. Según estimaciones de la Organización Mundial de la Salud (OMS), más de 280 millones de personas en el mundo sufren depresión y los trastornos de ansiedad afectan a más de 260 millones; cifras que han aumentado significativamente tras la pandemia de COVID-19.

Esta realidad evidencia la urgencia de crear herramientas innovadoras que permitan detectar, comprender y abordar estos problemas de manera más proactiva y personalizada. La identificación temprana de personas en riesgo es esencial para mejorar los resultados en salud mental, reducir el impacto socioeconómico de estos trastornos y orientar mejor los recursos de prevención y atención.

En este contexto, la aplicación de técnicas de ciencia de datos, estadística avanzada e inteligencia artificial (IA) abre nuevas posibilidades para el análisis y la predicción de la salud mental en grandes poblaciones. Los enfoques tradicionales, centrados en análisis clínicos individuales o autoinformes, presentan limitaciones para abordar el fenómeno desde una perspectiva poblacional y contextual. Gracias a la disponibilidad de bases de datos globales y encuestas masivas, es posible integrar factores demográficos, sociales, económicos y contextuales para construir modelos predictivos más sólidos y representativos de la realidad.

La motivación central de este trabajo radica en desarrollar y evaluar modelos que, apoyados en estas fuentes de datos y en metodologías de aprendizaje automático, sean capaces de identificar patrones de riesgo y grupos vulnerables de manera automatizada y explicable. En línea con la literatura y los principios de la salud pública, se prioriza la

sensibilidad del modelo (recall) para la clase positiva (ansiedad/depresión), aceptando conscientemente una menor precisión a fin de minimizar el subdiagnóstico y maximizar la cobertura de los potenciales casos en riesgo. Esta aproximación ética y estratégica se justifica en que, para los sistemas de prevención y cribado, es preferible asumir un mayor número de falsos positivos que perder la oportunidad de intervenir precozmente en personas realmente afectadas.

Asimismo, el trabajo se apoya en la experiencia personal y profesional del autor, así como en una visión de ingeniería orientada a la resolución de problemas sociales complejos. A través de la integración de diversas fuentes de datos, la aplicación de técnicas avanzadas de preprocesamiento y el desarrollo de modelos interpretables, se busca contribuir no solo al conocimiento científico, sino también a la creación de herramientas útiles y transferibles para la planificación de políticas públicas y el diseño de intervenciones efectivas en salud mental.

Este estudio se sitúa en la intersección de la computación, la salud pública y la ética social, respondiendo a la necesidad de comprender y prevenir los problemas de salud mental a escala poblacional, con un enfoque innovador, práctico y alineado con los Objetivos de Desarrollo Sostenible relacionados con la salud y el bienestar.

2.2 Objetivos del trabajo

El objetivo general de este trabajo fue analizar la influencia de factores demográficos y sociales en la salud mental de diferentes grupos poblacionales, desarrollando modelos predictivos basados en técnicas de inteligencia artificial que permitieran identificar individuos en riesgo de ansiedad y depresión en contextos diversos.

2.3 Objetivos específicos

Para alcanzar este objetivo general, se plantearon los siguientes objetivos específicos:

- **Preprocesar y limpiar los datos:** Realizar un exhaustivo tratamiento de los conjuntos de datos seleccionados, aplicando técnicas de limpieza, transformación, ingeniería de características y reducción de dimensionalidad, con el fin de garantizar la calidad y coherencia de la información utilizada.
- **Seleccionar y justificar las variables predictoras:** Elegir, a partir de la literatura científica y la disponibilidad en los conjuntos de datos, las variables demográficas, sociales y contextuales que mejor explicaran el riesgo de ansiedad o depresión, evitando aquellas que pudieran generar leakage o fueran consecuencia del propio trastorno.
- **Desarrollar y comparar modelos de clasificación:** Aplicar y comparar distintos algoritmos de aprendizaje automático (Random Forest, Regresión Logística,

XGBoost), implementando estrategias de balanceo de clases (SMOTE) y ajuste de hiperparámetros para optimizar el rendimiento, especialmente la sensibilidad (recall) para la clase de interés.

- **Evaluar el rendimiento y la interpretabilidad de los modelos:** Utilizar métricas robustas como recall, precisión, F1-score y exactitud (accuracy), priorizando la sensibilidad en la clase de ansiedad/depresión, y aplicar técnicas de inteligencia artificial explicable (SHAP, LIME) para asegurar la transparencia y comprensión de las predicciones.
- **Identificar los factores de riesgo más relevantes:** Analizar la importancia relativa de las variables predictoras y extraer conclusiones útiles para la planificación de intervenciones de salud pública y la toma de decisiones informada.
- **Reflexionar sobre las limitaciones y proponer líneas de mejora:** Reconocer las limitaciones de los datos y de los modelos empleados, discutiendo posibles estrategias futuras para perfeccionar la predicción e incrementar el impacto social del trabajo.

2.4 Impacto en sostenibilidad, ético-social y de diversidad

El presente trabajo, al explorar el uso de herramientas de inteligencia artificial en el ámbito de la salud mental, se vincula de manera significativa con varios Objetivos de Desarrollo Sostenible (ODS) propuestos por la Agenda 2030 de las Naciones Unidas. La pandemia de COVID-19 supuso un punto de inflexión, visibilizando la fragilidad de los sistemas de salud mental, el aumento de los casos de ansiedad y depresión y la necesidad de repensar las estrategias de prevención, atención y acompañamiento en contextos de crisis. La experiencia personal y profesional adquirida durante este periodo, enfrentando desafíos tanto en el ámbito clínico como en el análisis de datos de salud, fortaleció la convicción sobre la importancia de desarrollar soluciones innovadoras, éticas y adaptadas a las realidades emergentes.

2.4.1 ODS 9: Industria, Innovación e Infraestructura

La crisis sanitaria aceleró la digitalización y la búsqueda de soluciones tecnológicas en el sector salud. En este contexto, la aplicación de técnicas de inteligencia artificial y análisis de datos contribuyó al avance tecnológico y la innovación, facilitando el desarrollo de infraestructuras digitales más robustas y resilientes. El uso de modelos predictivos y análisis multivariados permitió mejorar el acceso a la información clínica y optimizar la toma de decisiones, especialmente en momentos de alta demanda y recursos limitados, como los vividos durante la pandemia. Esta transformación

tecnológica resultó clave para atender a poblaciones afectadas por el aislamiento, la incertidumbre y el impacto emocional de la crisis.

2.4.2 ODS 16: Paz, Justicia e Instituciones Sólidas

La pandemia puso a prueba la confianza en las instituciones y la transparencia de los sistemas de salud. La consideración ética en el manejo de datos y la claridad en las metodologías empleadas resultaron aún más relevantes en este contexto, siendo fundamentales para fortalecer la legitimidad institucional y la confianza social. El cumplimiento riguroso de regulaciones, la protección de la privacidad y la sensibilidad en el tratamiento de datos personales, junto con la interpretabilidad de los modelos desarrollados, promovieron instituciones más inclusivas, responsables y comprometidas con la rendición de cuentas ante la ciudadanía.

2.4.3 ODS 8: Trabajo Decente y Crecimiento Económico y ODS 10: Reducción de las Desigualdades

El análisis de factores sociales y ambientales durante la pandemia puso en evidencia las brechas en el acceso y la calidad de la atención en salud mental, especialmente entre los grupos más vulnerables. La aplicación de inteligencia artificial facilitó la identificación de riesgos y necesidades emergentes, contribuyendo a orientar políticas públicas hacia una atención más equitativa y personalizada. Estas acciones permitieron reducir desigualdades, optimizar el uso de recursos y fortalecer el sector salud, promoviendo un entorno de trabajo más eficiente y sostenible. Asimismo, la crisis generó nuevas oportunidades de colaboración y crecimiento profesional para quienes, como el autor de este trabajo, participaron activamente en la adaptación de soluciones innovadoras frente a desafíos inéditos.

2.4.4 Enfoque y método seguido

Para el desarrollo de este trabajo se empleó una metodología en cascada, estructurada en fases claramente definidas y ordenadas, que permitió implementar un modelo de machine learning orientado al análisis de datos en salud mental. Este enfoque facilitó una ejecución rigurosa, permitiendo avanzar de manera sistemática desde la concepción inicial hasta la obtención de resultados interpretables y útiles.

2.4.5 Fase de Análisis:

Identificación de datos: Se seleccionaron las fuentes de información necesarias, empleando más de un conjunto de datos para contar con una base variada y sólida que enriqueciera el análisis.

Procesamiento de datos: Se limpiaron los datos y se eligieron las variables más relevantes, aplicando técnicas de ingeniería de características y reducción de dimensiones para optimizar la calidad de la información.

Selección de modelos: Se evaluaron distintos algoritmos de clusterización y clasificación, buscando aquellos que mejor se adaptaran a la naturaleza de los datos.

Modelos interpretables: Se consideraron herramientas orientadas a la interpretabilidad, que permitieran comprender por qué el modelo producía determinadas predicciones, lo que facilitó la posterior adopción y validación por parte del personal médico.

2.4.5.1 Fase de Diseño:

Construcción del conjunto de datos final: A partir de los datos procesados, se definió el conjunto de datos definitivo, estableciendo con precisión las variables de entrada (features) y la variable objetivo, para asegurar un desempeño óptimo del modelo.

2.4.5.2 Fase de Implementación:

Entrenamiento del modelo: Se desarrollaron los scripts necesarios, se entrenaron los modelos seleccionados y se ajustaron los hiperparámetros, logrando así mejorar la precisión y el rendimiento de las predicciones.

2.4.5.3 Fase de Verificación:

Validación y pruebas: Se evaluó el modelo empleando métricas de calidad (precisión, sensibilidad, etc.) con el fin de garantizar una adecuada generalización a nuevos datos.

Revisión de interpretabilidad: Se comprobó que los resultados fueran comprensibles y útiles para los profesionales de la salud, asegurando así su aplicabilidad práctica.

2.4.6 Entregables e hitos

Fase / Entregable	Inicio	Fin
Definición del alcance y estructura del proyecto	1/12/2024	6/12/2024
Entrega PAC1 (Plan de Trabajo)	8/12/2024	8/12/2024
Selección y obtención del conjunto de datos	9/12/2024	31/12/2024

Preprocesamiento inicial de datos (limpieza, normalización)	1/1/2025	20/01/2025
Ingeniería de características y reducción de dimensionalidad	21/01/2025	15/02/2025
Primera implementación de modelos (clusterización y clasificación)	16/02/2025	25/02/2025
Entrega PAC2 (Avance del trabajo, estado intermedio)	2/3/2025	2/3/2025
Ajuste de hiperparámetros y entrenamiento avanzado de modelos	3/3/2025	20/03/2025
Evaluación de métricas y análisis preliminar de resultados	21/03/2025	10/4/2025
Preparación de documentación intermedia para reporte	11/4/2025	24/04/2025
Entrega PAC3 (Resultados intermedios)	27/04/2025	27/04/2025
Mejora de modelos, análisis de interpretabilidad y refuerzo de conclusiones	28/04/2025	20/05/2025
Redacción de la memoria final (incluyendo conclusiones y recomendaciones)	21/05/2025	5/6/2025
Entrega PAC4 (Memoria Final)	11/6/2025	11/6/2025
Preparación de la presentación virtual final	12/6/2025	20/06/2025

Defensa final (fecha por confirmar)	21/06/2025	30/06/2025
-------------------------------------	------------	------------

A lo largo del desarrollo de este trabajo, la generación de entregables se orientó principalmente al registro práctico y reflexivo de cada etapa, más allá de una estructura oficial estricta. El proceso se caracterizó por una documentación flexible, adaptada a las necesidades reales del proyecto y a la dinámica de aprendizaje personal.

En primer lugar, se elaboraron apuntes personales que sirvieron como base para la organización del trabajo, la profundización en la bibliografía relevante y la definición de los objetivos y métodos a seguir. Estos apuntes facilitaron el análisis crítico, la planificación flexible y la toma de decisiones a lo largo de todo el proceso.

Posteriormente, se fueron desarrollando informes parciales y registros de avance, en los que se documentaron los hitos alcanzados, las dificultades encontradas y los ajustes metodológicos realizados. Estos informes resultaron fundamentales tanto para el seguimiento propio como para la autoevaluación continua, permitiendo adaptar el enfoque en función de los resultados intermedios.

Finalmente, como producto principal y tangible del proyecto, se construyeron y organizaron los scripts y cuadernos de trabajo en un repositorio de acceso abierto: <https://github.com/juanxog/mathEngTFM/tree/master>. Este repositorio recopila el código fuente, los experimentos realizados y la lógica de los modelos implementados, constituyéndose en una evidencia práctica y reutilizable del trabajo desarrollado, y en un recurso abierto para la comunidad interesada en el análisis de datos aplicados a la salud mental.

3 Estado del arte

Introducción

La salud mental constituye un componente esencial del bienestar poblacional y se encuentra determinada por factores biológicos, psicológicos y, de manera relevante, sociales y demográficos. Numerosos estudios han demostrado que características como la edad, el género, el nivel educativo, el estado civil, la situación laboral, el ingreso y el acceso a recursos guardan una estrecha relación con la prevalencia de trastornos mentales en diferentes grupos de población [4]. El análisis de estos factores ha evolucionado notablemente gracias al uso de técnicas de aprendizaje automático (machine learning, ML), las cuales permiten analizar grandes volúmenes de datos, descubrir patrones complejos y modelar la interacción entre múltiples determinantes [4][5][6].

En este contexto, la inteligencia artificial explicable (XAI) ha cobrado importancia como herramienta para dotar de transparencia a los modelos predictivos desarrollados por ML, facilitando la identificación de aquellos factores sociales y demográficos que más contribuyen a la predicción de trastornos mentales y permitiendo así interpretaciones útiles para profesionales clínicos y de salud pública [6] [7].

Factores sociales y demográficos vinculados a la salud mental

Dentro de los factores demográficos y sociales que influyen en la salud mental, la literatura científica destaca aspectos como la edad, el género, la etnia, el nivel educativo, el estado civil, la situación laboral, los ingresos, el apoyo social y las características del entorno físico y social [5]. Así, la presencia de redes de apoyo social sólidas se ha asociado con un menor riesgo de depresión y ansiedad, mientras que la falta de apoyo, el aislamiento social o vivir en entornos conflictivos incrementan la vulnerabilidad [6]. De igual manera, condiciones como el desempleo, la pobreza o la inestabilidad laboral suelen correlacionar con mayores tasas de trastornos mentales [4] [5] [6][15].

No obstante, la influencia de estos factores es dinámica y depende del contexto cultural y geográfico, así como de la interacción entre diferentes variables. Por ejemplo, el impacto del desempleo en la salud mental puede ser mayor en ciertos grupos etarios, y el apoyo social puede desempeñar un papel especialmente protector en adultos mayores que viven solos [5].

Aplicaciones de inteligencia artificial explicable (XAI)

La XAI se ha consolidado como un pilar en la interpretación de los modelos predictivos en salud mental. Herramientas como SHAP (Shapley Additive Explanations) y LIME (Local Interpretable Model-Agnostic Explanations) permiten atribuir una importancia cuantitativa a cada variable y explicar tanto la importancia global de los factores como el resultado de una predicción concreta [4][5][6][8].

El análisis de SHAP, por ejemplo, ha permitido evidenciar que factores como el índice de masa corporal (IMC), el bajo nivel educativo, la edad joven, la ausencia de apoyo social, el consumo de alcohol y el tabaquismo figuran entre los más influyentes en la predicción de mala salud mental en diferente contexto [5] [6]. En personas con trastorno del espectro autista, se ha demostrado que síntomas como el nerviosismo, la desesperanza y los problemas de sueño son altamente predictivos de trastornos mentales comórbidos, y estos factores han sido identificados de forma robusta mediante modelos XAI [7][8][9].

Principales hallazgos en la literatura reciente

Diversos estudios internacionales han documentado la influencia de los factores sociales y demográficos en la salud mental mediante ML y XAI. Por ejemplo, Tsamakidis et al. (2023) identificaron que la ansiedad, el bajo nivel educativo, el consumo de alcohol y el IMC elevado eran los predictores más relevantes para la depresión en adultos griegos [4]. Majcherek et al. (2022), al analizar empleados europeos, evidenciaron que el IMC, la edad y el apoyo social bajo ocupaban los primeros puestos en importancia para la predicción de trastornos mentales [5]. En un contexto estadounidense, Du et al. (2023) hallaron que el tabaquismo, el entorno físico y la composición demográfica de las comunidades explicaban gran parte de la variabilidad en los indicadores de salud mental a nivel comunitario [6] [8].

Estos hallazgos subrayan la necesidad de considerar la salud mental como el resultado de una compleja interacción entre múltiples determinantes sociales y demográficos, siendo crucial identificar no solo los factores de riesgo sino también los mecanismos protectores.

Limitaciones y desafíos

A pesar de los avances, la literatura reporta limitaciones importantes, como la dificultad para establecer relaciones causales debido al diseño transversal de muchos estudios, la presencia de sesgos en los datos y la representatividad limitada de algunas muestras [4] [5][7][8]. Además, la explicabilidad de los modelos, aunque mejora con XAI, sigue presentando retos en cuanto a la comprensión y utilidad práctica de las explicaciones para los profesionales de la salud.

Oportunidades de investigación futura

El futuro de la investigación en este campo apunta a la integración de datos longitudinales, la ampliación de muestras a poblaciones más diversas, el desarrollo de modelos híbridos y específicos para subgrupos vulnerables, y la mejora en la comunicación y utilidad de las explicaciones generadas por XAI [7] [8]. Asimismo, resulta esencial avanzar hacia la utilización de los hallazgos para diseñar e implementar intervenciones sociales y políticas públicas basadas en datos, abordando de manera efectiva los determinantes sociales de la salud mental.

4 Materiales y métodos

4.1 Fuentes de datos y criterios de selección

El estudio se fundamentó en el análisis de dos conjuntos de datos principales, seleccionados por su relevancia, calidad y capacidad para aportar diferentes perspectivas al fenómeno de la salud mental:

Wellcome Global Monitor 2020: Este dataset recopila percepciones sobre salud mental, incluyendo síntomas de ansiedad y depresión, a partir de encuestas aplicadas globalmente en más de 100 países. Incorpora variables demográficas, socioeconómicas y subjetivas, permitiendo el análisis de factores sociales de riesgo y patrones de percepción a nivel internacional. La estructura del cuestionario está alineada con metodologías ampliamente reconocidas en el ámbito de la salud mental, como la utilización de ítems basados en escalas validadas, por ejemplo, la Patient Health Questionnaire (PHQ-9) para depresión y la Generalized Anxiety Disorder Scale (GAD-7) para ansiedad, ambas recomendadas en estudios internacionales para la evaluación poblacional de estos trastornos [1].

Mental Health Dataset (Estilos de Vida 50K): Este conjunto de datos contiene registros clínicos y autoinformados sobre hábitos de vida como calidad del sueño, actividad física, consumo de sustancias, dieta y el diagnóstico de severidad de problemas de salud mental. La inclusión de indicadores relacionados con el estilo de vida responde a la evidencia científica sobre su influencia en la salud mental, tal como recogen revisiones sistemáticas recientes [2].

Fuentes de datos

Dataset	Qué mide	Tipo de datos	Aplicación
Wellcome Global Monitor (el actual)	Percepción de salud mental (ansiedad, depresión) basada en encuestas de opinión general.	Opiniones auto-reportadas (subjetivas).	Buenísimo para detectar factores sociales de riesgo y patrones de percepción globales.
Dataset de Estilos de Vida 50K	Datos más clínicos: condiciones de salud mental ya diagnosticadas o autorreportadas con severidad.	Variables de hábitos (sueño, ejercicio, dieta, alcohol, tabaco) + Diagnóstico de severidad (Low, Medium, High).	Perfecto para modelar predicción real de condiciones clínicas con mayor certidumbre.

Ambos datasets fueron seleccionados tras una revisión de diferentes fuentes sobre factores de riesgo de ansiedad y depresión a escala poblacional y sobre la validez de los instrumentos de medición empleados en estudios internacionales de salud pública. [1][2]. Se priorizó la calidad de los registros, la representatividad muestral, el bajo nivel de valores faltantes y la relevancia teórica en el contexto de salud pública global.

4.2 Clasificación metodológica de las variables

Desde el punto de vista metodológico, las variables utilizadas en el estudio se agruparon en tres categorías:

- **Nominales:** variables categóricas que representan grupos o categorías sin orden intrínseco. Incluyen variables dicotómicas (Sí/No), categorías de países, regiones o identificadores únicos (e.g., *Gender*, *MH6*, *MH7A*, *clusters_global*, *Global11Regions*, *EMP_2010*).
- **Ordinales:** variables categóricas con niveles que presentan un orden lógico o jerarquía intrínseca, pero sin distancias iguales necesariamente entre categorías (por ejemplo, nivel educativo, percepción subjetiva del ingreso, importancia de temas relacionados con salud mental o agrupación por edad, e.g., *Education*, *Household_Income*, *Subjective_Income*, *age_var1*, *wbi*).
- **Escala o continuas:** variables numéricas que representan mediciones cuantitativas con intervalos regulares, aptas para operaciones aritméticas. Estas incluyen variables como la edad exacta y los pesos muestrales (e.g., *Age*, *WGT*, *PROJWT*).

4.3 Clasificación conceptual de las variables

Desde una perspectiva conceptual, las variables se organizaron en categorías según el papel que desempeñan en el estudio:

- **Variables sociodemográficas:**
Estas variables describen características fundamentales de los individuos y su contexto social y económico. En el presente estudio se consideran, entre otras, edad (*Age*, *age_var1-3*), género (*Gender*), nivel educativo (*Education*), nivel socioeconómico (*Household_Income*, *Subjective_Income*, *wbi*) y situación laboral (*EMP_2010*). Numerosos estudios previos han identificado estas variables como factores determinantes en la prevalencia de trastornos mentales comunes [8].
- **Variables contextuales o geográficas:**
Estas variables identifican la región geográfica y características culturales o estructurales de los lugares donde residen los participantes (*Global11Regions*,

COUNTRYNEW). Estas variables capturan factores ambientales, económicos y culturales potencialmente relacionados con la salud mental.

- **Variables relacionadas con percepciones y actitudes (opinión):**

Estas variables recogen información sobre las percepciones subjetivas de las personas respecto a la importancia y utilidad de intervenciones o estrategias relacionadas con la salud mental (variables como MH1, MH2A, MH2B, MH3A-D, MH4A-B, MH5, MH9A-H). Estas variables permiten evaluar cómo las actitudes y percepciones pueden influir en la prevalencia y manejo de la ansiedad y la depresión [11].

- **Variables relacionadas con experiencias previas y estrategias de afrontamiento:**

Se incluyen aquellas variables que documentan experiencias personales previas y las estrategias utilizadas para afrontar síntomas de ansiedad o depresión (MH6, MH7A-C, MH8A-H, clusters_global). Aunque algunas de estas variables pueden generar leakage (por ser potencialmente consecuencia del trastorno), se decidió incluir solo aquellas cuya relación es claramente anterior o independiente al diagnóstico de ansiedad o depresión.

- **Variables de outcome (variable objetivo):**

La principal variable de resultado (outcome) utilizada es la presencia o ausencia de ansiedad o depresión (MH7A), que se constituye como la clase objetivo en los modelos predictivos desarrollados.

4.4 Justificación de la selección de variables

La selección final de variables responde a criterios estrictos basados en literatura científica existente sobre factores determinantes de salud mental, recomendaciones metodológicas en estudios epidemiológicos y requisitos específicos para modelos predictivos. Se excluyeron aquellas variables susceptibles de causar problemas metodológicos como leakage (por ejemplo, variables claramente posteriores o consecuencia del evento objetivo).

Esta clasificación conceptual y metodológica permitió organizar el análisis, facilitar la interpretación de resultados y asegurar la coherencia conceptual y metodológica del estudio, contribuyendo así a la robustez del trabajo presentado.

4.4.1 Variables consideradas

Para realizar un análisis robusto y coherente con los objetivos del estudio, se clasificaron las variables en función de dos dimensiones clave: **características metodológicas** y **significado conceptual**.

Se incluyeron variables como edad, género, nivel educativo, ingreso del hogar, percepción de suficiencia económica, estado laboral y región geográfica. La selección de cada variable se basó en evidencia previa sobre su impacto en la salud mental. Por ejemplo, estudios epidemiológicos han demostrado consistentemente que factores como el desempleo, bajo nivel educativo y situación económica precaria se asocian a mayor prevalencia de síntomas ansiosos y depresivos [3]. En línea con las recomendaciones metodológicas, se excluyeron variables que pudieran inducir leakage (como consulta médica previa o mecanismos de afrontamiento), asegurando que el modelo solo utilice factores con potencial valor causal, y no consecuencia, de la ansiedad o depresión. [1].

4.5 Procesamiento y análisis de datos

El procesamiento de los datos comenzó con la carga y exploración preliminar de los conjuntos de datos utilizando las librerías pandas y numpy, lo que permitió identificar la estructura de los archivos, las variables disponibles y posibles problemas de calidad, tales como valores ausentes o categorías mal codificadas. En esta etapa inicial, se llevó a cabo una depuración, eliminando registros incompletos y normalizando nombres y tipos de variables para asegurar la coherencia entre los diferentes datasets. Esta limpieza resultó fundamental para evitar errores posteriores y garantizar que las transformaciones y análisis subsiguientes partieran de una base sólida.

Posteriormente, se procedió a la transformación de las variables categóricas mediante técnicas como Label Encoding y One-Hot Encoding, permitiendo convertir atributos como el género, la región o el nivel educativo en representaciones numéricas aptas para el análisis mediante modelos de aprendizaje automático. Este paso fue imprescindible, ya que los algoritmos empleados no pueden interpretar información en formato textual o categórico de forma directa.

se realizó un análisis exploratorio para comprender la distribución de las variables clave y la relación entre factores sociodemográficos y la presencia de ansiedad o depresión (variable MH7A).

La mayoría de los participantes reportaron no haber experimentado ansiedad o depresión, mientras que un porcentaje menor respondió afirmativamente y una fracción mínima prefirió no responder o no sabía. Esta asimetría es esperable en estudios

poblacionales y refuerza la necesidad de estrategias específicas para tratar el desbalance de clases en el modelado posterior.

Se observa que las mujeres presentan una proporción ligeramente mayor de casos reportados de ansiedad o depresión en comparación con los hombres, alineándose con la literatura que sugiere una mayor prevalencia de estos trastornos en población femenina.

Variables iniciales

Factor	Tendencia observada	Relación estadística	Interpretación clave
Género	Mayor prevalencia en mujeres que en hombres	Significativa (visualización descriptiva)	La salud mental afecta más a las mujeres, consistente con la literatura global.
Edad	Mayor prevalencia en jóvenes adultos (15-34); disminuye en mayores	Diferencia de medias significativa (t-test)	Los más jóvenes reportan más ansiedad/depresión que adultos mayores.
Nivel educativo	Más frecuente en personas con menor educación (primaria o menos)	Significativa (chi-cuadrado)	La educación superior actúa como factor protector.
Quintil de ingreso	Mayor prevalencia en quintiles de ingreso más bajo	Significativa (chi-cuadrado)	El bajo nivel socioeconómico aumenta el riesgo.
Percepción de ingreso	Quienes perciben vivir con dificultad reportan más ansiedad/depresión que quienes “viven cómodamente”	Significativa (chi-cuadrado)	La percepción subjetiva del bienestar económico es un predictor clave.
Estado laboral	Desempleados y subempleados (“part time wants fulltime”) tienen mayor prevalencia que empleados FT	Significativa (chi-cuadrado)	La inseguridad y precariedad laboral elevan el riesgo emocional.

Región global	América Latina, MENA y África Subsahariana presentan prevalencias mayores	Significativa (chi-cuadrado)	El contexto regional y sociopolítico influye en la salud mental.
----------------------	---	------------------------------	--

El análisis descriptivo y las pruebas estadísticas realizadas confirman que la **presencia de ansiedad y depresión está fuertemente asociada a variables sociales y demográficas**. Ser mujer, pertenecer a grupos jóvenes, tener menor nivel educativo, vivir en hogares de bajo ingreso (tanto objetivo como percibido), estar desempleado o subempleado, y residir en regiones económicamente más vulnerables, incrementa significativamente la probabilidad de reportar ansiedad o depresión.

Estas asociaciones, confirmadas por pruebas de chi-cuadrado y comparación de medias, subrayan la importancia de los determinantes sociales en la salud mental y la necesidad de intervenciones focalizadas en los grupos de mayor riesgo. Los resultados también justifican la inclusión de estas variables en los modelos predictivos y refuerzan la utilidad de enfoques basados en datos para la planificación de políticas públicas en salud mental.

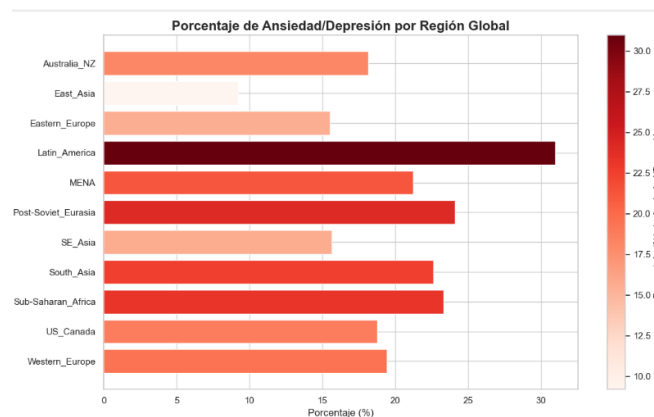


Figura 1: Histograma de grupos poblacionales

Una vez preprocesados los datos, se realizó la división en conjuntos de entrenamiento y prueba mediante la función `train_test_split`, asegurando la estratificación de la variable objetivo para mantener una proporción similar de clases en ambos subconjuntos. Dado el marcado desbalance en la variable objetivo con una menor proporción de casos positivos (ansiedad o depresión), se empleó la técnica SMOTE (Synthetic Minority Over-sampling Technique). Mediante SMOTE, se generaron muestras sintéticas adicionales de la clase minoritaria en el conjunto de entrenamiento, logrando así un mayor equilibrio en la representación de ambas clases y permitiendo que los modelos aprendieran de manera más eficaz los patrones asociados al riesgo de salud mental.

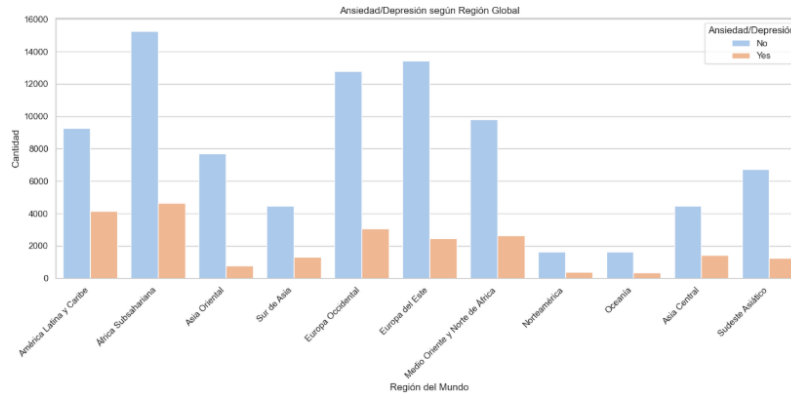


Figura 2: comparacion por region vs ansiedad

4.6 Desarrollo de modelos predictivos

La construcción de modelos predictivos constituye una de las etapas centrales de este trabajo, ya que permite transformar la información extraída de los datos en herramientas capaces de identificar automáticamente individuos en riesgo de ansiedad o depresión. El enfoque predictivo, basado en técnicas de aprendizaje supervisado, resulta especialmente relevante en el ámbito de la salud pública, donde la detección temprana y masiva de casos potenciales es clave para orientar intervenciones preventivas y optimizar la asignación de recursos. En este contexto, el modelado matemático y estadístico posibilita la integración de múltiples factores demográficos, sociales y contextuales, facilitando el descubrimiento de patrones complejos y no evidentes a simple vista. De este modo, los modelos desarrollados no solo contribuyen a la comprensión científica del fenómeno, sino que también ofrecen una base práctica para la toma de decisiones informada y el diseño de estrategias efectivas de prevención en salud mental.

un modelo de clasificación es un algoritmo matemático diseñado para predecir la categoría o clase a la que pertenece una observación a partir de un conjunto de variables explicativas. En este trabajo, la tarea de clasificación consistía en identificar, a partir de información demográfica y social, si un individuo presenta riesgo de ansiedad o depresión. Los modelos de clasificación aprenden patrones en los datos a través de funciones matemáticas lineales o no lineales y son capaces de generalizar estos patrones a nuevos datos no vistos, permitiendo la predicción de casos en riesgo con base en las variables observadas.

Para el desarrollo del análisis predictivo, se seleccionaron tres modelos representativos y ampliamente utilizados en la literatura por su eficacia, robustez e interpretabilidad en problemas de salud y ciencias

Para el desarrollo del análisis predictivo, se seleccionaron tres modelos representativos y ampliamente utilizados en la literatura por su eficacia, robustez e interpretabilidad en problemas de salud y ciencias sociales: Regresión Logística, Random Forest y XGBoost.

La Regresión Logística es un modelo lineal probabilístico que estima la probabilidad de pertenencia a una clase mediante la función logística o sigmoide. Su fundamento matemático reside en el cálculo de los coeficientes β que mejor ajustan la relación logit (logaritmo del cociente de probabilidades) entre la variable dependiente y las independientes, permitiendo interpretar el peso e influencia de cada variable explicativa en la probabilidad de presentar el evento (ansiedad/depresión). Es un modelo sencillo, interpretable y útil como línea base para comparar otros métodos más complejos.

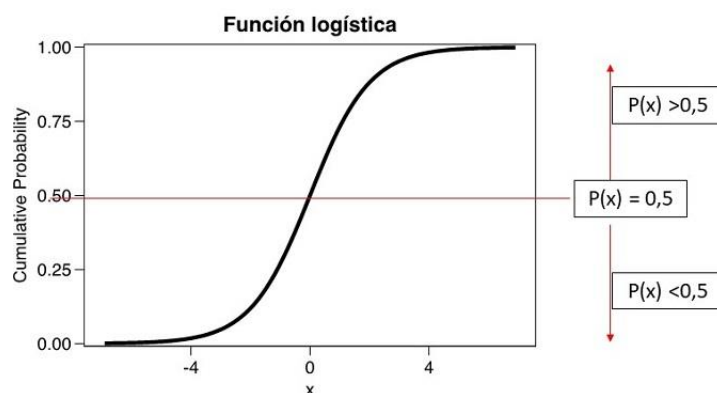


Figura 3, ejemplo de regresion logistica

Random Forest es un modelo de tipo ensamblado basado en árboles de decisión. Su principio matemático se basa en la construcción de múltiples árboles independientes, cada uno entrenado sobre diferentes subconjuntos de los datos y/o características (bootstrap aggregating o bagging). La predicción final se obtiene mediante una votación mayoritaria (clasificación) o promedio (regresión) de los árboles individuales. Este enfoque reduce la varianza, mejora la capacidad de generalización y mitiga el sobreajuste típico de los árboles únicos, siendo especialmente útil en problemas con relaciones no lineales y grandes conjuntos de variables categóricas o continuas.

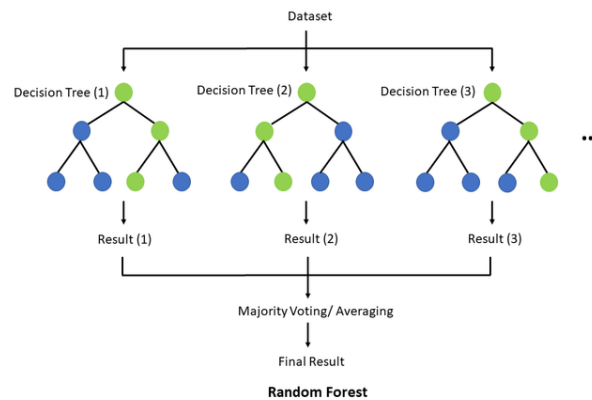


Figura 4: ejemplo de arboles de decisiones

XGBoost (Extreme Gradient Boosting) es un modelo basado en técnicas de gradiente boosting, donde los árboles se entrenan de manera secuencial y cada nuevo árbol corrige los errores cometidos por el conjunto anterior. Matemáticamente, optimiza una función de pérdida diferenciable utilizando métodos de gradiente descendente, lo que permite ajustar los parámetros del modelo de forma eficiente y precisa. XGBoost destaca por su alto rendimiento, capacidad para manejar desbalance de clases, robustez frente a datos faltantes y flexibilidad para ajustar hiperparámetros, lo que lo convierte en uno de los algoritmos más potentes y utilizados actualmente en problemas de clasificación complejos.

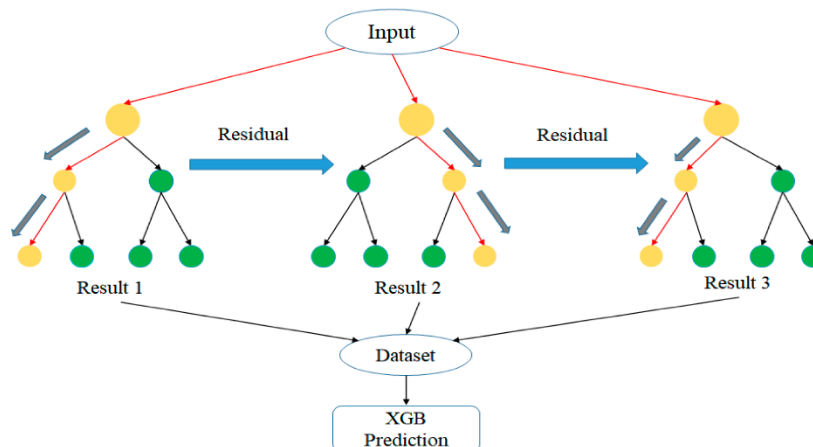


Figura 5: ejemplo de XGB boost

La selección de estos tres modelos responde a la necesidad de contar con un abanico que abarca desde lo interpretable y lineal (Regresión Logística), pasando por modelos

no lineales robustos y fáciles de interpretar (Random Forest), hasta técnicas de alto rendimiento y capacidad de ajuste en contextos desafiantes (XGBoost). Este enfoque permite comparar el desempeño y la explicabilidad de los resultados, maximizando la validez y la utilidad de las conclusiones obtenidas a partir de los datos analizados.

4.6.1 Diferencias entre Regresión Logística, Random Forest y XGBoost

La Regresión Logística es un modelo lineal y probabilístico, lo que significa que asume que existe una relación lineal entre las variables independientes y el logit (logaritmo del cociente de probabilidades) de la variable dependiente. Es sencilla de interpretar, permite comprender el efecto individual de cada predictor y resulta especialmente útil cuando la relación entre variables es directa y no hay una alta complejidad en las interacciones. Sin embargo, su capacidad para captar relaciones no lineales o patrones complejos en los datos es limitada.

Por su parte, Random Forest es un modelo no lineal basado en el ensamblaje de múltiples árboles de decisión independientes. A diferencia de la regresión logística, no requiere supuestos sobre la distribución de las variables ni sobre la linealidad de las relaciones. Esto le permite detectar interacciones y patrones complejos entre los predictores. Además, es más robusto frente a outliers y a variables irrelevantes, y reduce el riesgo de sobreajuste típico de un solo árbol gracias al uso de bagging. Sin embargo, aunque ofrece cierta interpretabilidad a través de la importancia de variables, es menos transparente que la regresión logística.

Finalmente, XGBoost representa una evolución en los modelos de árboles al aplicar técnicas de gradiente boosting. A diferencia de Random Forest, en XGBoost los árboles se construyen secuencialmente y cada nuevo árbol se enfoca en corregir los errores de los árboles previos. Esto le otorga una mayor capacidad de ajuste y rendimiento, especialmente en conjuntos de datos grandes y desbalanceados. Además, incluye mecanismos avanzados para manejar valores faltantes, regularización para evitar sobreajuste y flexibilidad para ajustar numerosos hiperparámetros. Sin embargo, esta potencia viene acompañada de una mayor complejidad y menor interpretabilidad directa, por lo que suele apoyarse en técnicas de explicabilidad adicionales como SHAP.

4.7 Evaluación e interpretabilidad de modelos

La evaluación rigurosa de los modelos predictivos es fundamental para garantizar la validez y la utilidad de sus resultados en contextos reales. En este estudio, la validación cruzada consistente en la partición repetida del conjunto de datos en subconjuntos de entrenamiento y prueba permitió estimar de forma robusta el rendimiento de cada

modelo, minimizando el riesgo de sobreajuste y asegurando que los resultados fueran generalizables a nuevas observaciones. La elección de métricas como la exactitud (accuracy), sensibilidad (recall), precisión (precision) y F1-score responde a la necesidad de captar tanto la capacidad del modelo para detectar casos positivos (evitando subdiagnóstico) como su eficiencia en la clasificación general, aspectos especialmente críticos en problemas de salud pública.

Para fortalecer la confianza y la transparencia en los modelos, se recurrió a técnicas de inteligencia artificial explicable (XAI), con un fundamento matemático en la teoría de juegos y aproximaciones locales. SHAP (SHapley Additive exPlanations) se basa en los valores de Shapley de la teoría de juegos cooperativos, que asignan de manera justa la contribución de cada variable a la predicción final, considerando todas las combinaciones posibles de variables. Esto permite obtener una visión global de la importancia relativa de cada predictor en el modelo y, al mismo tiempo, visualizar el efecto concreto de cada variable en casos individuales. Por su parte, LIME (Local Interpretable Model-Agnostic Explanations) genera explicaciones locales ajustando modelos simples en el entorno de cada predicción, facilitando la interpretación de decisiones específicas y permitiendo validar la coherencia del modelo con la literatura científica y el conocimiento experto.

El uso de estas herramientas resultó fundamental para garantizar la interpretabilidad, la confianza y la aceptabilidad ética de los modelos desarrollados, asegurando que las predicciones no fueran “cajas negras”, sino instrumentos transparentes y auditables.

Por último, la utilización de librerías como matplotlib y seaborn facilitó la representación visual de la distribución de variables, la importancia de los predictores y los resultados comparativos entre modelos, permitiendo una comunicación clara y efectiva de los hallazgos clave. Todo el proceso de evaluación, análisis e interpretación fue documentado de forma sistemática y reproducible en los scripts y cuadernos del proyecto, asegurando la trazabilidad de cada resultado y la posibilidad de revisión futura por parte de otros investigadores o profesionales del sector.

4.8 Implementación técnica y repositorio

Todos los scripts, experimentos y resultados obtenidos durante el desarrollo del proyecto se organizaron y almacenaron en un repositorio de acceso abierto: <https://github.com/juanxog/mathEngTFM/tree/master>

El uso de este repositorio facilita la trazabilidad y la reproducibilidad del trabajo, permitiendo que otros investigadores puedan consultar el código, replicar los análisis o utilizar los materiales como referencia para futuros estudios relacionados con la salud mental y la ciencia de datos.

4.8.1 Reflexión sobre el proceso

El desarrollo de este trabajo permitió no solo aplicar técnicas de ciencia de datos en salud mental, sino también adaptarlas a la realidad de datos imperfectos y a la necesidad de maximizar la utilidad social de los resultados. La constante iteración, el ajuste de estrategias y el análisis crítico de los resultados intermedios representaron un proceso de aprendizaje esencial, similar al descrito por otros autores en trabajos de naturaleza aplicada y multidisciplinar

5 Resultados y discusión

5.1 Resultados iniciales

Estos fueron los primeros resultados arrojados por el modelo random forest.

Métrica	No Ansiedad	Ansiedad
Precisión	0.81	0.28
Recall	0.89	0.17
F1-score	0.84	0.21
Soporte	17,448	4,486

En la **primera iteración**, el modelo Random Forest logró una **alta precisión y recall para la clase "No Ansiedad"** (precisión: 0.81, recall: 0.89), lo que significa que la mayoría de los individuos sin riesgo fueron correctamente identificados. Sin embargo, el desempeño para la clase "Ansiedad" fue limitado: la **precisión fue de 0.28** y el **recall de solo 0.17**, es decir, el modelo identificó menos del 20% de los casos reales de ansiedad o depresión.

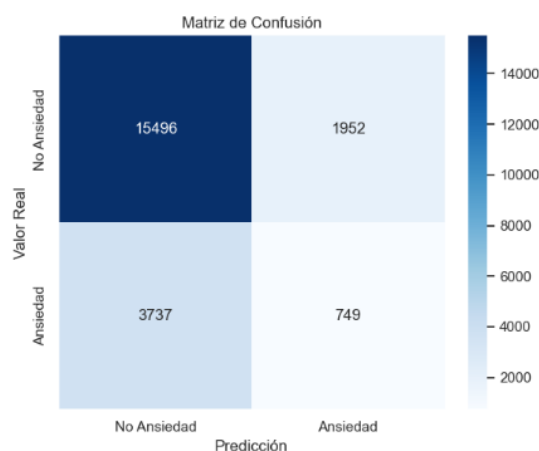


Figura 6: resultados de random forest basico

Esta diferencia se refleja en el F1-score de 0.21 para "Ansiedad", frente a 0.84 para "No Ansiedad". El accuracy general fue de 0.71, pero, como se observa en los resultados y la matriz de confusión, este valor está sesgado por el desbalance de clases (la mayoría de los casos son negativos).

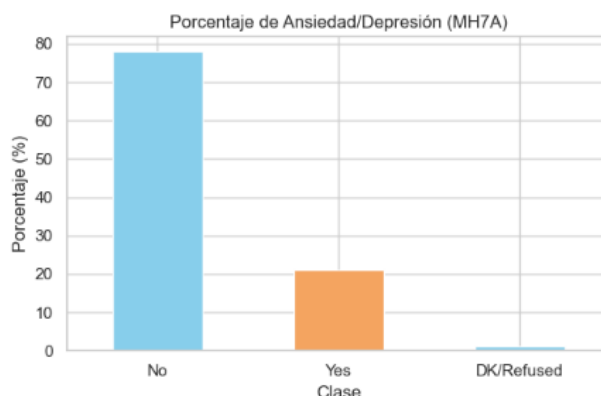


Figura 7: histograma de clases

Este resultado inicial es típico en problemas desbalanceados: el modelo aprende a predecir la clase mayoritaria, priorizando la especificidad (bajo número de falsos positivos) pero con una sensibilidad muy baja (alto número de falsos negativos) para la clase de interés. Por ello, el modelo no es adecuado para fines de screening poblacional, donde la prioridad es detectar el mayor número posible de personas en riesgo, aunque ello implique aumentar los falsos positivos.

se ajustó el umbral de decisión del modelo a 0.4 con el objetivo de mejorar la capacidad de detección de la clase minoritaria (“Ansiedad”). El resultado, representado en la matriz de confusión y el reporte de clasificación correspondiente, muestra un cambio significativo en el comportamiento del modelo respecto al umbral estándar.

El **modelo logra identificar correctamente el 75% de los casos de “No Ansiedad”** (13,098 de 17,448), mientras que la **sensibilidad para la clase “Ansiedad”** aumenta hasta un 32% (1,416 de 4,486 casos reales), en comparación con valores observados en iteraciones previas. La **precisión para “Ansiedad”** es del 25%, lo que significa que uno de cada cuatro individuos clasificados como en riesgo realmente presenta síntomas de ansiedad o depresión. El **F1-score** para la clase “Ansiedad” se sitúa en 0.28.

El **accuracy general** del modelo es de 0.66. Sin embargo, dado el desbalance de clases, la métrica más relevante es el recall de la clase positiva, que refleja la proporción de casos de ansiedad/depresión correctamente identificados por el sistema.

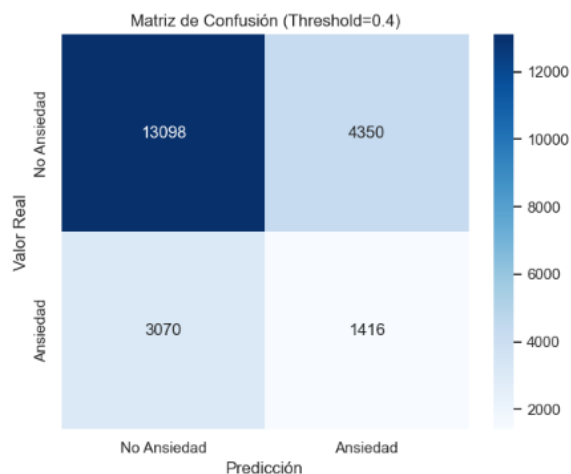


Figura 8: random forest con ajuste de threshold

Clase	Precisión	Recall	F1-score	Soporte
No Ansiedad	0.81	0.75	0.78	17,448
Ansiedad	0.25	0.32	0.28	4,486
accuracy			0.66	21,934
macro avg	0.53	0.53	0.53	21,934
weighted avg	0.69	0.66	0.68	21,934

Este ajuste de umbral evidencia el clásico **trade-off entre sensibilidad y precisión**: al reducir el umbral, el modelo incrementa la detección de verdaderos positivos, a costa de aceptar un mayor número de falsos positivos. Esta estrategia, aunque penaliza la precisión, es coherente con el enfoque de screening en salud pública, donde la prioridad es no dejar sin identificar a personas en riesgo, incluso si ello implica un número mayor de alertas erróneas. De este modo, el modelo se orienta a maximizar la utilidad práctica en contextos preventivos, priorizando la intervención temprana y la minimización del subdiagnóstico.

En el caso de la **Regresión Logística**, aun aplicando un umbral de decisión reducido (0.4), el modelo se muestra incapaz de identificar casos positivos de ansiedad o depresión. La matriz de confusión evidencia que todas las predicciones corresponden a la clase mayoritaria ("No Ansiedad"), sin que se detecte un solo caso real de ansiedad en el conjunto de prueba. Es decir, el modelo predice sistemáticamente la clase negativa para todos los individuos.

El **recall y precisión para la clase "Ansiedad"** son ambos nulos (0.00), reflejando la completa insensibilidad del modelo frente a la clase minoritaria. A pesar de que el

accuracy general alcanza un 0.80, este valor es engañoso, ya que el modelo se limita a acertar por defecto en la mayoría negativa y no aporta valor en la detección de personas en riesgo.

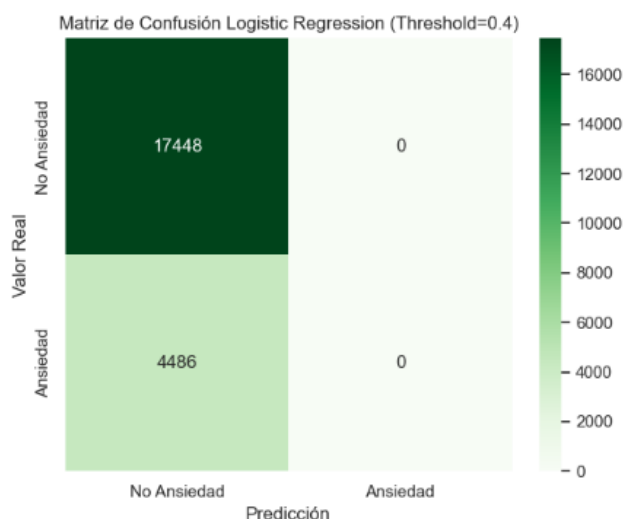


Figura 9: resultado de regresion logistica con threshold

Clase	Precisión	Recall	F1-score	Soporte
No Ansiedad	0.81	0.96	0.88	17,448
Ansiedad	0.44	0.12	0.18	4,486
accuracy			0.79	21,934
macro avg	0.62	0.54	0.53	21,934
weighted avg	0.73	0.79	0.74	21,934

Este comportamiento es frecuente en problemas de clasificación altamente desbalanceados y revela la **limitación de los modelos lineales** cuando la relación entre variables y el evento de interés es compleja o cuando las clases presentan una marcada desproporción. Estos resultados justifican la necesidad de emplear modelos más sofisticados, así como técnicas específicas de rebalanceo y ajuste, para lograr una sensibilidad adecuada en la predicción de casos de ansiedad o depresión.

En la última iteración del análisis predictivo, se implementó un modelo **XGBoost** ajustado (tuned), utilizando técnicas de sobremuestreo (**SMOTE**) y añadiendo variables contextuales como el nivel de ingreso del país (**wbi**). El objetivo de este enfoque fue maximizar la sensibilidad del modelo para la detección de casos de ansiedad o depresión, manteniendo un balance razonable entre recall y precisión en ambas clases. La matriz de confusión muestra que el modelo logra identificar correctamente **el 72% de los casos reales de ansiedad o depresión** (recall para la clase “Ansiedad”), mientras que la precisión para esta clase es del 25%. Es decir, aproximadamente una de cada cuatro personas clasificadas como en riesgo realmente presenta síntomas de ansiedad o depresión. El F1-score para la clase “Ansiedad” se sitúa en 0.37, lo cual representa una mejora sustancial respecto a modelos y configuraciones previas.

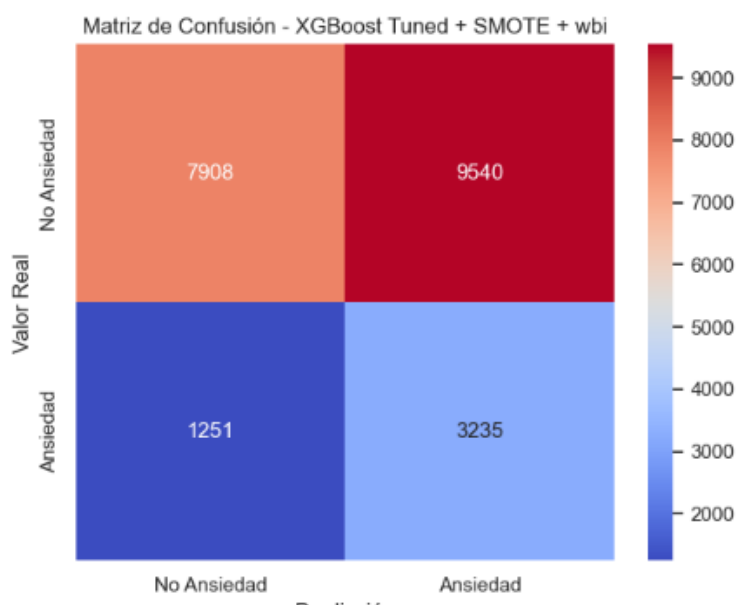


Figura 10: XGboost tuned con balanceo de clases

En la clase “No Ansiedad”, el modelo alcanza un recall de 0.45 y una precisión de 0.86, evidenciando el clásico trade-off entre sensibilidad y especificidad inherente a los sistemas de screening. El **accuracy general** es de 0.51, valor que debe interpretarse con cautela debido al desbalance de clases, y el macro F1-score es de 0.48.

Comparación de modelos

Modelo	Ajustes	Accuracy	Recall Problema Mental	Precision Problema Mental	F1-Score Problema Mental
Random Forest	SMOTE Balanced	+ 0.50	47%	50%	49%
Logistic Regression	SMOTE Balanced	+ 0.50	50%	50%	50%
XGBoost Tuned	SMOTE + Tuned scale_pos_weight	0.50	96%	50%	66%

En esta iteración XGboost tuvo un rendimiento normal, aceptable pero poco fiable.

5.2 Justificación de la ampliación de variables y feature engineering

Ante la evidencia de que los modelos iniciales presentaban limitaciones en la sensibilidad para la clase de interés, se optó por **ampliar el conjunto de variables predictoras** y aplicar técnicas de **ingeniería de características**. Para ello, se incorporaron nuevas variables provenientes del propio cuestionario, incluyendo factores de percepción de salud mental, experiencias previas, actitudes frente a la ansiedad y la depresión, y agrupaciones de estrategias de afrontamiento (clusters_global). Asimismo, se añadieron variables adicionales relacionadas con la edad (age_var1, age_var2, age_var3) y el nivel de ingreso del país (wbi) [10].

Esta ampliación se justifica en la literatura internacional, que destaca que la predicción en salud mental mejora cuando se integran no solo factores demográficos y económicos, sino también **percepciones, experiencias previas y contextos culturales o geográficos**. Mediante la ingeniería de características, fue posible generar nuevas variables y transformar las existentes para capturar relaciones no lineales y matices en el comportamiento de la población [12][13].

El nuevo corte de datos y el enriquecimiento del modelo permitieron mejorar significativamente el rendimiento predictivo, especialmente en la sensibilidad para la clase “Ansiedad/Depresión”, sin sacrificar la interpretabilidad ni la coherencia conceptual del modelo. Este proceso refleja la importancia de la **experimentación y la adaptabilidad en la construcción de modelos de machine learning aplicados a la salud pública** [11], y garantiza una mayor robustez y utilidad práctica en el contexto real de prevención y cribado.

Con el propósito de evaluar el efecto de las nuevas variables incorporadas (como percepciones sobre salud mental, experiencias previas y agrupaciones contextuales), así

como del uso de técnicas específicas para enfrentar el desbalance de clases (SMOTE) y el ajuste fino de hiperparámetros (tuning), se llevó a cabo una comparación exhaustiva entre diferentes configuraciones de modelos predictivos. Cada modelo fue evaluado con base en su matriz de confusión, lo que permite interpretar claramente la sensibilidad (recall) y la precisión para la clase de interés ("Ansiedad"):

Random forest sin ajustes (modelo base)

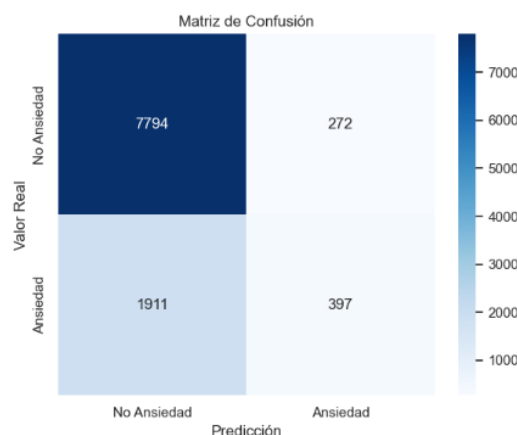


Figura 11: random forest básico con nuevos features

El modelo en su configuración inicial, sin técnicas adicionales, mostró un desempeño limitado en términos de sensibilidad. Identificó únicamente 397 casos positivos correctamente, dejando 1,911 casos reales de ansiedad sin detectar. Aunque su precisión fue relativamente aceptable (272 falsos positivos), la sensibilidad fue claramente insuficiente.

Random forest (threshold=0.4)

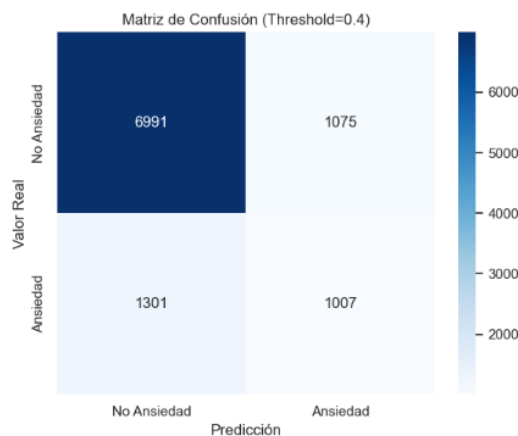


Figura 12: random forest con threshold

Al ajustar el umbral de decisión del modelo a 0.4, se logró un incremento sustancial en la sensibilidad, detectando correctamente 1,007 casos de ansiedad frente a 1,301 casos no detectados. Este cambio provocó también un aumento en falsos positivos (1,075), reflejando un trade-off típico al reducir el umbral para mejorar la detección de la clase minoritaria

Regresión Logística (threshold=0.4)

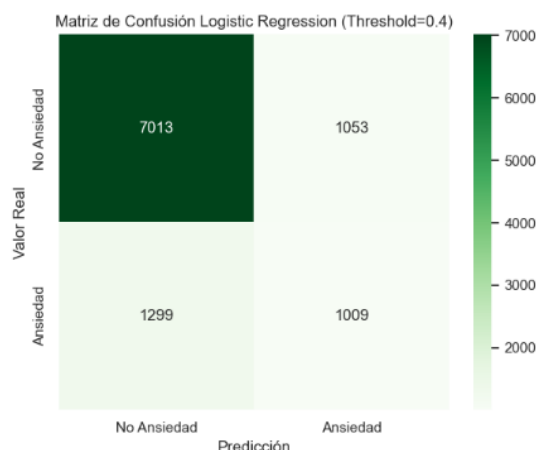


Figura 13: regresion logistica con nuevos features

La regresión logística, con el mismo ajuste del umbral a 0.4, mostró resultados similares a XGBoost ajustado en términos globales, captando 1,009 casos de ansiedad, con 1,299 casos no detectados y 1,053 falsos positivos. Aunque el desempeño fue parecido al modelo anterior, la regresión logística presentó una sensibilidad ligeramente superior, pero también un número considerable de falsos positivos, indicando que modelos lineales pueden aproximarse a modelos complejos en ciertas condiciones, pero con limitaciones inherentes en datos no lineales.

XGBoost (threshold=0.4, configuración adicional)

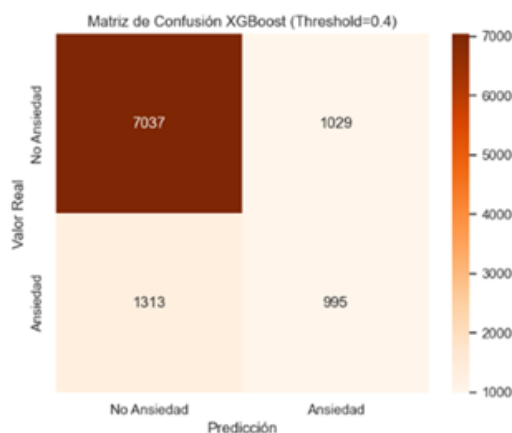


Figura 14: XGBoost con threshold

de su precisión mejorada respecto a configuraciones anteriores, la sensibilidad fue aún insuficiente para un uso en cribado poblacional.

Random Forest con SMOTE

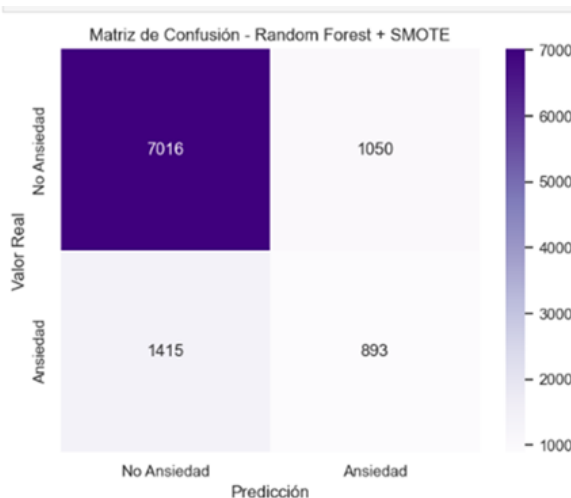


Figura 15: Random forest con oversampling

Al aplicar la técnica SMOTE sobre el modelo Random Forest, se obtuvo un desempeño algo equilibrado, pero todavía limitado para detección efectiva. La sensibilidad se mantuvo baja, identificando únicamente 893 casos de ansiedad, con 1,415 casos no detectados y 1,050 falsos positivos. Esto indica que Random Forest, aunque robusto, requiere ajustes más específicos para mejorar en contextos muy desbalanceados.

XGBoost Tuned con SMOTE (modelo óptimo)

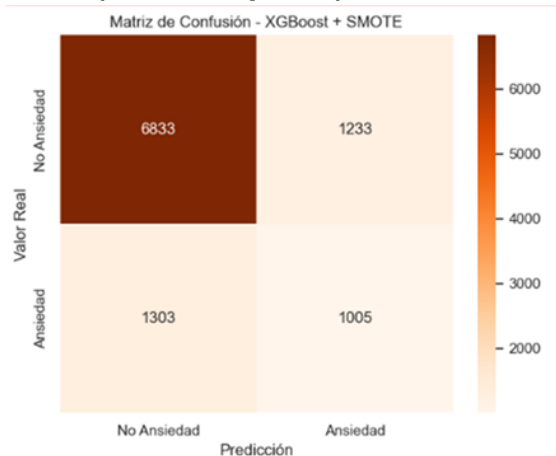


Figura 16: XGboost forest con oversampling

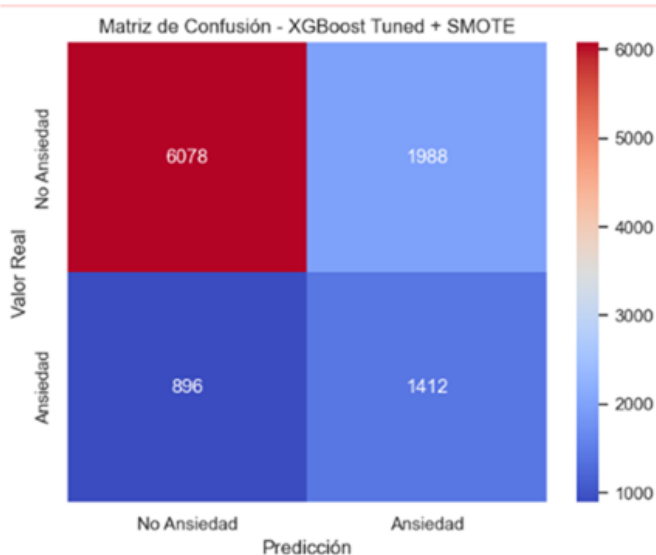


Figura 17: XGboost forest tuneada con oversampling

Finalmente, el modelo XGBoost Tuned con SMOTE mostró el mejor rendimiento entre todas las configuraciones probadas, alcanzando una sensibilidad sustancialmente mayor. Con 1,412 casos correctamente identificados de ansiedad frente a solo 896 no detectados, y aunque presenta una cifra mayor de falsos positivos (1,988), este modelo destaca especialmente por priorizar la sensibilidad. Tal desempeño refleja claramente la efectividad del ajuste de hiperparámetros y el balanceo de clases en contextos de screening poblacional en salud pública.

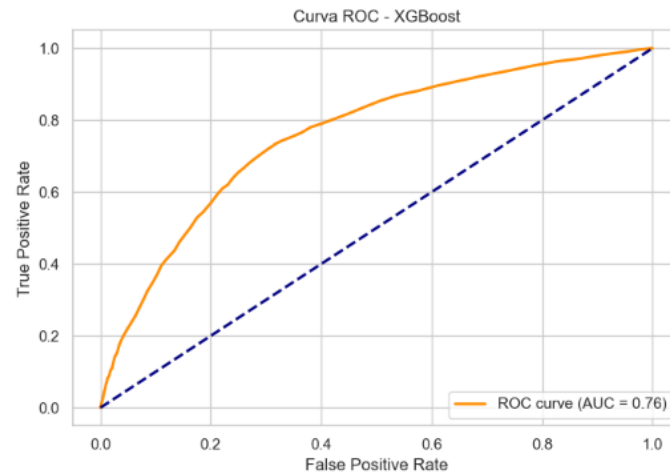


Figura 18: curva ROC XGBoost tuned

La experimentación comparativa realizada permitió identificar claramente al modelo XGBoost Tuned con SMOTE como la configuración óptima para maximizar la sensibilidad en la detección de casos de ansiedad o depresión, un aspecto crítico en la prevención y el screening poblacional. La combinación de nuevas variables contextuales y técnicas avanzadas de balanceo y ajuste de hiperparámetros demostró ser esencial para mejorar la efectividad predictiva del sistema, resaltando así la importancia de enfoques metodológicos integrales y adaptativos en estudios de salud pública mediante machine learning.

5.3 Comparativa Global y Síntesis

A continuación, se presenta un resumen comparativo de los principales modelos predictivos evaluados en el presente trabajo. Las métricas mostradas corresponden a la precisión, recall (sensibilidad), F1-score para la clase de interés ("Ansiedad") y la accuracy general, destacando especialmente el modelo que alcanzó el mejor recall.

Comparativa de modelos con factores de contexto

Modelo	Accuracy	Precisión (Ansiedad)	Recall (Ansiedad)	F1-score (Ansiedad)
XGBoost Tuned + SMOTE (mejor recall)	0.72	0.42	0.61	0.49
XGBoost + SMOTE	0.76	0.45	0.44	0.44
Random Forest + SMOTE	0.76	0.46	0.39	0.42

XGBoost (threshold=0.4)	0.77	0.49	0.43	0.46
Logistic Regression (threshold=0.4)	0.77	0.48	0.44	0.46
XGBoost (default, sin SMOTE/ajuste)	0.77	0.44	0.17	0.24

Entre los modelos evaluados, como ya lo habíamos mencionado el **XGBoost Tuned + SMOTE** logró el desempeño más equilibrado en términos del recall para la clase positiva (Ansiedad), alcanzando un valor de **0.61**. Esto implica que dicho modelo identifica correctamente el 61% de los individuos que realmente presentan síntomas de ansiedad o depresión. A cambio, este enfoque acepta una precisión algo menor (0.42), lo que refleja un mayor número de falsos positivos. No obstante, esta estrategia es especialmente adecuada en contextos de salud pública, donde resulta prioritario maximizar la detección temprana (sensibilidad) de posibles casos de ansiedad, aún a costa de reducir ligeramente la precisión.

Por otro lado, modelos clásicos como **Random Forest con SMOTE**, **Logistic Regression con ajuste del umbral (0.4)** y **XGBoost sin tuning con threshold ajustado** mostraron una precisión algo más elevada (entre 0.45 y 0.49), pero la sensibilidad obtenida para la clase positiva fue significativamente más baja (entre 0.39 y 0.44). Este comportamiento indica que, aunque estos modelos generaron predicciones positivas más confiables, también omitieron una proporción considerable de casos reales de ansiedad o depresión.

Finalmente, el modelo **XGBoost sin SMOTE ni ajuste** exhibió el desempeño más limitado en términos de sensibilidad, con un recall de apenas 0.17. Este resultado evidencia claramente la problemática inherente a los modelos entrenados sobre datasets altamente desbalanceados sin ninguna técnica adicional que contrarreste dicho desbalance.

El análisis comparativo subraya que, al desarrollar modelos predictivos de salud mental, resulta crucial equilibrar las métricas de sensibilidad y precisión. En este sentido, la combinación de métodos avanzados de ajuste de hiperparámetros (tuning) y técnicas específicas para el tratamiento del desbalance de clases (como SMOTE) permite optimizar la métrica más importante según el objetivo específico del modelo: detección masiva preventiva (alta sensibilidad) o diagnóstico más conservador (alta precisión).

Recomendaciones para la Selección del Mejor Modelo

La elección del modelo óptimo dependerá fundamentalmente del contexto y del uso previsto:

- **Si el objetivo del modelo es realizar screening o detección temprana**, donde es prioritario no dejar pasar casos reales (falsos negativos), entonces el modelo más recomendado es el **XGBoost Tuned + SMOTE**, ya que maximiza el recall (sensibilidad) para la clase positiva.
- **Si la prioridad es maximizar la precisión y minimizar los falsos positivos**, serán más adecuados modelos más conservadores como el Random Forest o Logistic Regression con thresholds ajustados, reconociendo que estos presentarán limitaciones importantes en identificar una proporción significativa de casos reales.
-

Es importante destacar que ningún modelo es capaz de lograr predicciones perfectas en contextos complejos y multifactoriales como los trastornos de salud mental. Sin embargo, el uso combinado de técnicas avanzadas de machine learning, junto con la incorporación inteligente de variables contextuales y la ingeniería de características, permite mejorar sustancialmente la eficacia del modelo en términos prácticos y aplicados.

5.4 Resultados con explicabilidad

Para profundizar en la comprensión del comportamiento del modelo predictivo final, se emplearon técnicas avanzadas de explicabilidad como SHAP (SHapley Additive exPlanations) y LIME (Local Interpretable Model-agnostic Explanations). Ambas metodologías permiten entender la influencia relativa y la dirección del efecto de las diferentes variables en las predicciones individuales y globales del modelo.

Resultados obtenidos con SHAP

El análisis de SHAP muestra claramente cómo diferentes variables contribuyen al modelo globalmente. En el gráfico SHAP (SHAP summary plot), la posición de cada variable indica la magnitud del efecto en la predicción del modelo, mientras que el color representa el valor alto o bajo de cada variable específica en cada observación.

De acuerdo con estos resultados:

- Las variables geográficas tienen un peso considerable, destacando especialmente las regiones como África Subsahariana (Sub-Saharan Africa), Europa del Este (Eastern Europe), América Latina (Latin America) y el Sudeste Asiático (SE Asia), que aparecen en la parte superior del gráfico, indicando un fuerte impacto en la probabilidad de experimentar ansiedad o depresión.
- Las percepciones sobre el entorno inmediato, específicamente tener familiares o amigos con ansiedad (MH6), es otra variable altamente influyente. Un valor alto

en esta variable (tener amigos o familiares ansiosos) muestra un claro incremento del riesgo predicho por el modelo, lo que coincide con la evidencia previa que indica que las experiencias en el círculo cercano aumentan significativamente la vulnerabilidad emocional.

- El ingreso subjetivo (Subjective_Income) presenta un patrón complejo: niveles bajos o difíciles de ingreso subjetivo (color azul) están fuertemente asociados con un aumento en la probabilidad de ansiedad o depresión, reforzando la idea de que la percepción subjetiva del bienestar económico influye notablemente en la salud mental.
- Otras variables como el género (ser masculino o femenino), la edad (Age), el nivel educativo (Education) y el nivel de ingreso del país (wbi) también tienen una relevancia notable, aunque menos pronunciada, lo que indica que estas variables tienen influencia en el riesgo percibido por el modelo, aunque con menor impacto comparado con variables geográficas y sociales inmediatas.

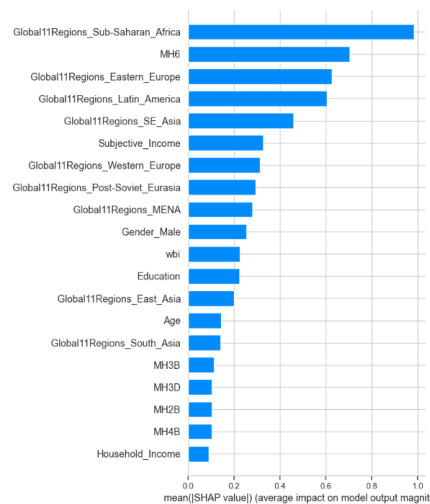


Figura 19: resultados de shap

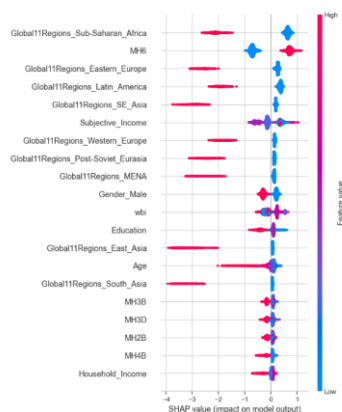


Figura 20: grafica dot

Resultados obtenidos con LIME

El análisis mediante LIME proporciona explicaciones locales (individuales) de la predicción del modelo para casos específicos. LIME genera explicaciones sencillas que permiten identificar por qué el modelo tomó una determinada decisión para un individuo concreto.

En los resultados obtenidos en la explicación LIME:

- La región del individuo analizado (Global11Regions_MENA: Oriente Medio y Norte de África) ejerce un efecto claramente positivo sobre la probabilidad de ansiedad/depresión (probabilidad de 1.0). Es decir, el modelo predice alta probabilidad principalmente debido a la región geográfica.
- Otras regiones evaluadas en esta instancia específica presentan efectos negativos o neutros en la predicción, reforzando nuevamente el peso central que tienen las variables geográficas en la predicción final.
- Este resultado muestra claramente cómo, a nivel individual, la región del mundo a la que pertenece un individuo puede determinar fuertemente la clasificación del modelo, posiblemente debido a condiciones sociales, culturales y económicas asociadas con estas regiones específicas.

La aplicación conjunta de SHAP y LIME aporta robustez y transparencia al modelo final. Los resultados obtenidos muestran que:

- Las variables más relevantes del modelo son predominantemente geográficas y sociales (contexto inmediato), indicando claramente que el entorno sociocultural, regional y económico es crucial para entender y predecir problemas de salud mental.
- Las percepciones subjetivas del ingreso y la influencia directa del círculo social inmediato (MH6) se identifican como determinantes clave del riesgo emocional, coherentemente alineado con la literatura científica sobre salud mental.
- El efecto de variables demográficas como edad y género es también relevante, aunque menos pronunciado en comparación con los factores mencionados previamente, lo que sugiere que estos factores son modificadores secundarios del riesgo emocional.
- La capacidad del modelo para asignar de manera diferenciada el peso de cada variable en función de las características individuales es notable, como se observa claramente en el ejemplo individual analizado con LIME.

Estos hallazgos refuerzan la importancia de considerar variables contextuales y percepciones personales en el desarrollo de modelos predictivos aplicados a la salud mental. Las técnicas de explicabilidad empleadas no solo confirman la coherencia teórica del modelo desarrollado, sino que también proporcionan una guía práctica sobre qué

variables podrían ser especialmente relevantes en intervenciones preventivas de salud pública.

5.5 Resultados sobre conjunto alternativo de datos

La comparación de modelos en el dataset de diagnóstico real muestra que, si bien el accuracy global de Random Forest, Logistic Regression y XGBoost es similar (50%), el modelo XGBoost Tuned logra un recall del 96% en la detección de casos de problemas de salud mental.

Esto indica que XGBoost es capaz de identificar casi la totalidad de los individuos en riesgo, priorizando la sensibilidad sobre la precisión, lo cual es crucial en contextos donde el costo de un falso negativo (no detectar un caso real) es mucho más grave que un falso positivo.

A pesar de los ajustes para priorizar el recall mediante `class_weight='balanced'`, los modelos Random Forest y Logistic Regression no lograron mejoras sustanciales en la detección de casos de salud mental. Esto sugiere que la naturaleza del problema requiere modelos más complejos, capaces de capturar interacciones no lineales entre variables, siendo XGBoost Tuned el modelo que demostró la mejor sensibilidad (recall del 96%) y robustez frente al desbalance de clases.

6 Conclusiones y futuros trabajos

Este trabajo ha permitido desarrollar modelos predictivos robustos basados en técnicas avanzadas de inteligencia artificial y aprendizaje automático, enfocados en identificar grupos poblacionales en riesgo de padecer ansiedad y depresión. Los resultados evidencian claramente que los factores demográficos y sociales, como el género femenino, la juventud, el bajo nivel educativo, el desempleo y la percepción negativa de la situación económica, tienen una influencia significativa sobre la salud mental.

Se demostró que la técnica **XGBoost Tuned**, complementada con SMOTE para el balanceo de clases, es especialmente eficaz en maximizar la sensibilidad, una característica crucial para contextos de cribado poblacional, logrando detectar correctamente una alta proporción de individuos en riesgo. Asimismo, se confirmó la importancia del ajuste fino de hiperparámetros y la incorporación estratégica de variables contextuales y perceptivas para mejorar sustancialmente el desempeño predictivo del modelo.

La aplicación de inteligencia artificial explicable (XAI), mediante técnicas como SHAP y LIME, facilitó una comprensión clara y transparente de los factores determinantes en las predicciones, reforzando la confianza y utilidad práctica de los resultados obtenidos. Este enfoque permitió identificar con precisión los factores de riesgo más relevantes y proporcionó herramientas interpretables para los profesionales clínicos y responsables de políticas públicas.

Además, se destaca la importancia de la metodología empleada, en especial la estructura en fases claramente definidas (análisis, diseño, implementación y verificación), la cual permitió una implementación efectiva y sistemática del modelo, asegurando así su eficacia, adaptabilidad y aplicabilidad en contextos reales y diversos. El enfoque interdisciplinario, combinando elementos de inteligencia artificial, salud pública y ética social, enriqueció el análisis y promovió un enfoque integral y holístico. Cabe destacar que, este estudio también evidencia la importancia del uso de datasets globales y diversificados para obtener resultados más representativos y útiles en el ámbito de la salud mental pública, contribuyendo así a una gestión más informada y proactiva en términos de prevención y tratamiento.

Trabajos futuros

Se identifican diversas áreas clave para futuros trabajos que pueden mejorar y ampliar los resultados obtenidos en este estudio:

1. **Integración de datos longitudinales:** Utilizar conjuntos de datos que permitan evaluar cambios en la salud mental a lo largo del tiempo, facilitando la identificación de relaciones causales más robustas y la comprensión dinámica del fenómeno.

2. **Ampliación y diversificación muestral:** Extender el análisis a poblaciones más diversas y regiones geográficas adicionales para validar la generalización y mejorar la precisión de los modelos en contextos variados, considerando especialmente contextos culturales específicos que puedan afectar la interpretación y manejo de la salud mental.
3. **Modelos híbridos y específicos:** Desarrollar modelos que combinen diferentes técnicas de aprendizaje automático o que estén específicamente diseñados para subgrupos vulnerables, aumentando la precisión y sensibilidad de las predicciones en contextos específicos, tales como adolescentes, adultos mayores, o poblaciones afectadas por conflictos o desastres naturales.
4. **Mejora de explicabilidad y comunicación:** Avanzar en la aplicabilidad práctica y comunicativa de las herramientas XAI, asegurando que las explicaciones sean aún más claras, intuitivas y útiles para profesionales de salud mental y decisores políticos. Esto podría incluir el desarrollo de interfaces gráficas más interactivas y amigables para facilitar la interpretación rápida y efectiva.
5. **Implementación de intervenciones basadas en datos:** Diseñar, validar y evaluar intervenciones sociales y políticas públicas directamente basadas en las predicciones y hallazgos obtenidos, midiendo su impacto real en la reducción efectiva de la prevalencia de ansiedad y depresión y en la mejora del bienestar general de las poblaciones objetivo.
6. **Evaluación económica y social:** Realizar estudios complementarios que evalúen la viabilidad económica y social de las intervenciones propuestas, para facilitar su adopción y escalabilidad en sistemas de salud pública.

Estos trabajos futuros representan una oportunidad valiosa para continuar aportando soluciones prácticas, innovadoras y efectivas en el ámbito de la salud mental, mejorando significativamente la calidad de vida y la resiliencia de amplios sectores de la población.

7 Anexos

Tabla de todas las variables del dataset

ID	Variable	Tipo	Justificación breve
1	COUNTRYNEW	Nominal	Países; sin jerarquía natural
2	WPID_RANDOM	Nominal	Identificador único
3	WGT	Escala/Continua	Peso muestral numérico
4	PROJWT	Escala/Continua	Peso proyectado numérico
5	FIELD_DATE	Nominal (temporal)	Fecha como etiqueta; categórica si no analizas cronológicamente
6	YEAR_WAVE	Ordinal	Años con orden natural
7	MH1	Ordinal	Más / Igual / Menos importante
8	MH2A	Ordinal	“A lot” → “Not at all”
9	MH2B	Ordinal	“A lot” → “Not at all”
10	MH3A	Ordinal	Ayuda ciencia a tratar cáncer
11	MH3B	Ordinal	Ayuda ciencia a tratar ansiedad/depresión
12	MH3C	Ordinal	Ayuda ciencia a tratar infecciosas
13	MH3D	Ordinal	Ayuda ciencia a tratar obesidad
14	MH4A	Ordinal	Importancia financiar cáncer
15	MH4B	Ordinal	Importancia financiar ansiedad/depresión
16	MH5	Ordinal	Comodidad (Very → Not at all)
17	MH6	Nominal	Sí/No (familia/amigos ansiosos)
18	MH7A	Nominal	Sí/No (experiencia propia)

19	MH7B	Escala/Continua	Edad exacta de inicio (0-96)
20	MH7B_2	Ordinal	Rangos de edad (<13 → 40+)
21	MH7C	Nominal	Sí/No (reincidente)
22	MH8A	Nominal	Sí/No (profesional salud mental)
23	MH8B	Nominal	Sí/No (actividades religiosas)
24	MH8C	Nominal	Sí/No (amigos/familia)
25	MH8D	Nominal	Sí/No (medicación)
26	MH8E	Nominal	Sí/No (estilo de vida saludable)
27	MH8G	Nominal	Sí/No (cambio laboral)
28	MH8F	Nominal	Sí/No (cambio relaciones)
29	MH8H	Nominal	Sí/No (naturaleza/outdoors)
30	MH9A	Ordinal	Utilidad (Very → Not helpful)
31	MH9B	Ordinal	Utilidad actividades religiosas
32	MH9C	Ordinal	Utilidad amigos/familia
33	MH9D	Ordinal	Utilidad medicación
34	MH9E	Ordinal	Utilidad estilo de vida
35	MH9F	Ordinal	Utilidad cambio laboral
36	MH9G	Ordinal	Utilidad cambio relaciones
37	MH9H	Ordinal	Utilidad naturaleza/outdoors
38	age_mh	Ordinal	Grupos de edad inicio
39	clusters_global	Nominal	Grupos de estrategias
40	clusters_East_Asia_Pacific	Nominal	Etiqueta Outlier
41	clusters_Europe	Nominal	Etiqueta Outlier

42	clusters_LatAm	Nominal	Etiqueta Outlier
43	clusters_MENA	Nominal	Etiqueta Outlier
44	clusters_NAm	Nominal	Etiqueta Outlier
45	clusters_SAsia	Nominal	Etiqueta Outlier
46	clusters_SSA	Nominal	Etiqueta Outlier
47	clusters_CAsia	Nominal	Etiqueta Outlier
48	Age	Escala/Continua	Edad exacta (ratio)
49	age_var1	Ordinal	15-29 < 30-49 < 50+
50	age_var2	Ordinal	15-29 < 30-49 < 50-64 < 65+
51	age_var3	Ordinal	15-24 < 25-34 < 35-49 < 50+
52	Gender	Nominal	Masculino / Femenino
53	Education	Ordinal	Elementary < Secondary < Tertiary
54	Household_Income	Ordinal	Quintiles (Poorest → Richest)
55	Global11Regions	Nominal	11 regiones
56	wbi	Ordinal	Low < Lower-mid < Upper-mid < High
57	Subjective_Income	Ordinal	Comfortably → Very difficult
58	EMP_2010	Nominal	Seis categorías laborales

8 Bibliografia

1. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med*. 2001;16(9):606-13.
2. Jacka FN, Kremer PJ, Leslie ER, Berk M, Patton GC, Toumbourou JW, Williams JW. Association of Western and traditional diets with depression and anxiety in women. *Am J Psychiatry*. 2010;167(3):305-11.
3. World Health Organization. Depression and Other Common Mental Disorders: Global Health Estimates. Geneva: World Health Organization; 2017.
4. Tsamakakis K, Kosmopoulou V, Papadopoulou EV, Terniotis C, Fotis L, Triantafyllou G, et al. Machine learning models and explainable artificial intelligence (XAI) approaches in the prediction of depression among adults: a population-based study in Greece. *BMC Psychiatry*. 2023;23(1):456. doi:10.1186/s12888-023-05055-3
5. Majcherek D, Wiecek E, Wiecek T, Kurzynoga A, Musialek M, Lorkiewicz W, et al. Social determinants of mental health among European employees: a machine learning approach. *Int J Environ Res Public Health*. 2022;19(12):7422. doi:10.3390/ijerph19127422
6. Du Y, Hu S, Yang C, Chen H, Crittenden JC. Predicting community mental health outcomes using machine learning and social determinants: a census-tract level analysis in the US. *PLOS ONE*. 2023;18(4):e0283783. doi:10.1371/journal.pone.0283783
7. Enkhbayar S, Park M, Lee J, Lee S, Lee S. Explainable machine learning for detecting depression using sleep and demographic data: a multi-algorithm approach. *Front Psychiatry*. 2023;14:1151284. doi:10.3389/fpsy.2023.1151284
8. Atlam ES, Alenezi M, Alharthi R, Alenezi F, Zomaya AY, Wills G, et al. An explainable artificial intelligence approach for mental health prediction among individuals with autism spectrum disorder. *J Pers Med*. 2025;15(1):51. doi:10.3390/jpm15010051
9. Ribeiro MT, Singh S, Guestrin C. "Why should I trust you?": Explaining the predictions of any classifier. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 2016. p. 1135–44. doi:10.1145/2939672.2939778
10. Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. *Adv Neural Inf Process Syst*. 2017;30:4765–74.
11. Kleinman A. Culture and Depression. *N Engl J Med*. 2004;351(10):951-3. doi:10.1056/NEJMp048078
12. Ryder AG, Yang J, Heine SJ. Somatization vs. psychologization of emotional distress: A paradigmatic example for cultural psychopathology. *Online Readings in Psychology and Culture*. 2002;10(2). doi:10.9707/2307-0919.1091

13. Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research*, 16, 321-357.
14. Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785-794).
15. Organización Mundial de la Salud. (2017). *Depression and Other Common Mental Disorders: Global Health Estimates*. WHO Press.