Salud, Ciencia y Tecnología. 2025; 5:1970
doi: 10.56294/saludcyt20251970

ORIGINAL

Random Forest modeling of bipolar affective disorder in Ecuador

Modelización mediante bosques aleatorios del trastorno afectivo bipolar en Ecuador

Cristhian Ismael Gómez Gaona[1], Andrea del Rocío Mejía Rubio[2], José Rubén León Pérez[3], Jesús Rodríguez[4], Zilma Diago Alfes[5], Laura Esther Muñoz Escobar[5], Cristian Inca[6], Jimmy Yaguana Torres[7], Marco Velasco Arellano[5]

[1]Universidad Estatal de Milagro, Ecuador.
[2]Universidad Técnica de Ambato, Ecuador.
[3]Universidad Autónoma de Tlaxcala. Campus Calpulalpan, México.
[4]Universidad UTE, Ecuador.
[5]Universidad Nacional de Chimborazo, Ecuador.
[6]Escuela Superior Politécnica de Chimborazo (ESPOCH), Ecuador.
[7]Clínica SurHospital, Ecuador.

Citar como: Gómez Gaona CI, Mejía Rubio A del R, León Pérez JR, Rodríguez J, Diago Alfes Z, Muñoz Escobar LE, et al. Random Forest modeling of bipolar affective disorder in Ecuador. Salud, Ciencia y Tecnología. 2025;5:1970. https://doi.org/10.56294/saludcyt20251970

Enviado: 09-02-2025 **Revisado:** 27-05-2025 **Aceptado:** 07-08-2025 **Publicado:** 08-08-2025

Editor: Prof. Dr. William Castillo-González

Autor para la correspondencia: Cristian Inca

ABSTRACT

Bipolar affective disorder is a mental disorder characterized by depressive and manic or hypomanic episodes. The complexity of the diagnosis of bipolar affective disorder due to the overlapping of its symptoms with other mood disorders led researchers and doctors to search for new and advanced techniques for the precise detection of bipolar affection disorder. One of these methods is the use of advanced machine learning algorithms under a statistical methodology for building logistical regression models, Random Forest. Support vector machines, Decision Tree, K-Nearest Neighbors, and Gradient Boosting, with 146 data collected from the psychiatric services affiliated with the mental health system of Ecuador. At the inferential level, the results suggest that the implementation of automatic algorithms based on the different methodologies for building models enables the successful prediction or classification of individuals with bipolar affective disorders in Ecuador compared to controlled patients who do not profile under this pathological picture. It is the best Random Forest statistical model (89,35 %) that dictates the best performance metrics compared to the Gradient Boosting model. The evolution of the overall prevalence of bipolar affective disorders in Ecuador over the past 22 years has increased by a small differential. However, from 2020 to 2022, there has been a considerable increase in the percentage prevalence of cases of bipolar affective disorders in Ecuador.

Keywords: Algorithms; Machine Learning; Decision Tree Gradient Boosting; Random Forest.

RESUMEN

El trastorno afectivo bipolar representa un trastorno mental caracterizado por episodios depresivos y maníacos o hipomaníacos. La complejidad en el diagnóstico del trastorno afectivo bipolar debido a la superposición de sus síntomas con otros trastornos del estado de ánimo llevó a los investigadores y médicos a buscar nuevas y avanzadas técnicas para la detección precisa del trastorno afectivo bipolar. Uno de estos métodos es el uso de algoritmos avanzados de aprendizaje automático bajo una metodología estadística para la construcción de modelos de Regresión logística, Random Forest. Máquinas de vectores de soporte, Decision Tree, K-Vecinos más cercanos, Gradient Boosting, con datos 146 recopilados de los servicios de psiquiatría adscritos al

© 2025; Los autores. Este es un artículo en acceso abierto, distribuido bajo los términos de una licencia Creative Commons (https://creativecommons.org/licenses/by/4.0) que permite el uso, distribución y reproducción en cualquier medio siempre que la obra original sea correctamente citada

sistema de salud mental del Ecuador. A nivel inferencial, los resultados sugieren que la implementación de algoritmo automáticos basado en las diferentes metodologías para la construcción de modelos permitan predecir o clasificar con éxito a individuos con trastornos afectivos bipolar en Ecuador en comparación con pacientes controles que no se perfilan bajo este cuadro patológico. Siendo el mejor modelo estadístico Random Forest (89,35 %) que dictamina las mejores métricas de rendimiento en comparación con el modelo Gradient Boosting. La evolución de la prevalencia general en trastornos afectivos bipolar en el Ecuador durante los últimos 22 años, se ha incrementado en un pequeño diferencial. Sin embargo, durante el año 2020 hasta 2022 se ha producido un incremento considerable en la prevalencia porcentual de casos en trastornos afectivos bipolar en el Ecuador.

Palabras clave: Algoritmos; Aprendizaje Automático; Decision Tree Gradient Boosting; Random Forest.

INTRODUCCIÓN

El trastorno bipolar (TB) es uno de los más graves, que afecta por igual a hombres y mujeres con estimaciones en padecimientos entre un 2 a un 3 % de la población a nivel mundial, según cifras de la Organización Mundial de la Salud (OMS),[1] lo que se traduce en más de 47 millones de personas en el mundo que combinan episodios depresivos con precedencia a síntomas depresivos superpuestos con el trastorno depresivo mayor (TDM) y periodos de episodios maníaco/hipomaníaco.

Este trastorno incluye euforia o irritabilidad, mayor actividad o energía, y otros síntomas como aumento de la verborrea, pensamientos acelerados, mayor autoestima, menor necesidad de dormir, distracción y comportamiento impulsivo e imprudente en la prevalencia clínica de los pacientes vulnerables esta enfermedad. Sin embargo, existen opciones de tratamiento efectivas que incluyen psicoeducación, reducción del estrés y fortalecimiento del funcionamiento social, y medicación.[2]

Las personas con trastorno bipolar han de tener como consecuencia la generación de algún tipo de discapacidad con una prevalencia en desarrollar un riesgo mayor de suicidio. Las tasas de suicidio entre los pacientes con TB son de 20 a 30 veces más altas que en la población general,[3] y el TB está relacionado con una muerte prematura significativa por diversas razones.[3] El diagnóstico de TB puede ser un desafío y puede requerir un largo proceso de diagnóstico que dure entre 6 a 10 años.[4]

Los individuos afectados frecuentemente presentan un episodio depresivo significativo y son diagnosticados erróneamente con depresión unipolar. Al diagnosticar erróneamente, lo que condiciona otorgar prescripciones inadecuadas de antidepresivos que no estabilizan el estado de ánimo y aumentan la probabilidad de episodios maníacos.[5] Por consiguiente, los trastornos bipolares no tratado se asocia con episodios de humor más prolongados y graves, así como con una mayor frecuencia de intentos de suicidio.[2]

Identificar a las personas en riesgo de trastorno bipolar podría permitir una evaluación dirigida, intervenciones tempranas y terapias más adecuadas. Un reciente análisis exhaustivo de ensayos de tratamiento para prevenir la enfermedad bipolar indicó una dependencia de antecedentes familiares para la identificación del riesgo.[3] Sin embargo, dada la naturaleza multifactorial del trastorno bipolar, la mayoría de las personas afectadas no tendrían antecedentes familiares favorables. Además, a diferencia de que se presenta en un cuadro de esquizofrenia que se caracteriza por trastornos significativos en la percepción de la realidad y cambios en el comportamiento, incluyendo ilusiones persistentes, alucinaciones, experiencias de influencia, pensamiento desorganizado, comportamiento altamente desorganizado, “síntomas negativos” y agitación extrema o ralentización de los movimientos.[4] En el contexto del trastorno bipolar, no se reconocen signos de alerta temprana.[5] Sería útil desarrollar nuevos métodos para identificar el riesgo que no dependan de los signos o síntomas clínicos existentes.

Las revisiones de investigaciones previas en el campo indican que existe un precedente notable para evaluar los registros médicos electrónicos (EHR) longitudinales. Esto, junto con el análisis predictivo, puede abrir nuevas vías para identificar factores de riesgo para la presentación de la enfermedad bipolar. Como se demostró anteriormente, estos datos pueden producir un fenotipado diagnóstico confiable de los pacientes bipolares.[6,7]

La práctica clínica ahora puede incluir hallazgos científicos, ya que el aprendizaje automático se utiliza ampliamente en la industria médica. Uno de los hallazgos científicos implica la creación de un modelo que permite la diferenciación a nivel individual de varias enfermedades y luego aplicaron el modelo a los datos recién recopilados para realizar clasificación o predicción. Una técnica de aprendizaje automático supervisado llamada Support Vector Machine (SVM) tuvo como objetivo detectar patrones de cambios cerebrales espacialmente dispersos en varios vóxeles.[8]

Otra de las contribuciones, se basó en la identificación temprana del riesgo de trastorno bipolar mediante el uso de modelos predictivos entrenados en diversas cohortes en los Estados Unidos que podría mejorar la evaluación específica de las personas de alto riesgo, reducir los diagnósticos erróneos y mejorar la asignación

de recursos limitados de salud mental. Se desarrollaron modelos predictivos válidos con múltiples algoritmos en cada sitio de estudio, incluyendo bosques aleatorios, máquinas de aumento de gradiente, regresión penalizada y algoritmos de aprendizaje de conjuntos apilados. Los predictores se basaron en datos ampliamente disponibles, independientes de un modelo de datos común que incluía datos demográficos, códigos de diagnóstico y medicamentos.

En consecuencia, el objetivo del presente estudio estuvo centrado en la descripción del fenómeno relacionado con trastorno afectivo bipolar en el Ecuador mediante la acogida de modelos de clasificación y predicción apropiados a los datos disponibles. Específicamente centrado para poder realizar una clasificación de pacientes con trastornos bipolar tipo I (TBI), bipolar tipo II (TBII) y depresión unipolar. Lo anterior, permitirá optar por un tratamiento acorde al diagnóstico sin cometer errores y además lograr rasgos característicos que permita un diagnóstico diferencial preciso e independiente de los estados de los trastornos afectivos.

MARCO TEÓRICO

Trastorno afectivo bipolar

Las fluctuaciones cíclicas y extremas del estado de ánimo, incluidos episodios de manía o hipomanía y tristeza, intercalados con intervalos de estado de ánimo estable, son el sello distintivo del trastorno bipolar, una enfermedad mental grave. La persona disfruta de euforia, mayor energía y actividad durante el período maníaco, mientras que la intensa negatividad y la falta de entusiasmo se encuentran entre los síntomas de la fase de depresión. Los cambios de humor pueden ocurrir con frecuencia y duración variables, según el tipo de enfermedad bipolar, la edad y otras circunstancias externas.[9]

Tipologías en trastorno afectivo bipolar

El trastorno bipolar (TB) es una afección grave del estado de ánimo caracterizada por episodios recurrentes de depresión e (hipo)manía. El trastorno bipolar se define por dos cambios de humor: altibajos emocionales (manía o hipomanía) y bajos (depresión). Aunque manía e hipomanía no son lo mismo, el Manual Diagnóstico y Estadístico de Trastornos Mentales DSM-5 divide el trastorno bipolar en tres subtipos principales: bipolar I, bipolar II y trastorno ciclotímico. Los dos subtipos principales son BD-I (episodios maníacos, típicamente acompañados de depresión) y BD-II (episodios hipomaníacos, combinados con depresión), mientras que el trastorno ciclotímico es una condición cíclica que produce períodos cortos de hipomanía y depresión.[10]

El trastorno unipolar es una enfermedad mental que perjudica significativamente el funcionamiento cotidiano y se caracteriza por un estado de ánimo persistentemente bajo o pérdida de interés en las actividades. Sin embargo, cuando el trastorno bipolar se identifica tempranamente, se puede prevenir o retrasar su inicio, y quienes ya lo padecen pueden tener mejores resultados clínicos.[11] En este sentido, la labor de diagnóstico es esencial, por consiguiente, se incluyeron cuatro procesos principales en el diagnóstico convencional del trastorno bipolar: sospecha basada en síntomas depresivos, seguimiento de la entrevista clínica y análisis de casos utilizando métodos de búsqueda basados en el diagnóstico y confirmación del diagnóstico.[12]

En este punto se centra la dificultad de diagnosticar la enfermedad bipolar, particularmente separarla de la depresión unipolar. El IHME,[13] propone protocolos para distinguir entre depresión mayor y trastorno bipolar, enfatizando que el trastorno bipolar se relaciona con una mayor labilidad del estado de ánimo, lentitud motora y duración del sueño, los cuales facilitan su identificación y con ellos las posibilidades de tratamiento.

Materiales

Se realizó recopilación de datos de ourworldindata.org, que describe la participación estimada de hombres y mujeres que presentaron trastorno bipolar en el Ecuador, en el caso de haber sido diagnosticados o no, basados en encuestas representativas, datos médicos y modelado estadístico.[14] Esta variable respuesta está representada por la prevalencia estimada asociada con la edad de individuos con trastorno bipolar, por cada 100 personas. Esta métrica se construyó siguiendo principios de comparabilidad mediante la estandarización por edad y registros en proporción de prevalencia con metodología uniforme con principio comparativo.

El Instituto Ecuatoriano de Seguridad Social (IESS) ofrece servicios de Psiquiatría con camas asignadas en cuatro hospitales en todo el país, donde también se proporciona atención ambulatoria. En las ciudades de Quito y Guayaquil, hay servicios de psiquiatría disponibles en consultorios ambulatorios. En Ecuador, hay cinco hospitales psiquiátricos en funcionamiento, tres en Quito, uno en Guayaquil y uno en Cuenca, con un total de 1 635 camas, lo que equivale a 12 camas por cada 100 000 habitantes. Los registros administrativos se han manejado con indicadores globales para proteger la privacidad médico-paciente.

El Instrumento de Evaluación de la Organización Mundial de la Salud para Sistemas de Salud Mental (OMS-IESM) se usó para recopilar información sobre el sistema de salud mental del Ecuador. En este sentido, se incluyeron en este estudio un total de 146 pacientes bajo monitoreo del departamento de psiquiatría adscrito al IESS, en los cuales se les evalúa diferentes métricas, tal como: sexo (1: mujer, 2: hombre), edad (Intervalos en grupos de edades), tipo de condición (1: bipolar II, 2: depresivo Unipolar, 3: bipolar I). En función con las condiciones

relacionadas, síntomas de melancolía (1: melancolía, 2: sin melancolía), hospitalización (1: hospitalizada, 2: ambulatoria), educación (educación agrupada en años), matrimonio (1: casado o concubino, 2: soltero), trabajo (1: paciente en actividad laboral o estudiando, 2: desempleado/pensionado por enfermedad). En este contexto, la detección del trastorno afectivo bipolar y posterior clasificación en sus tres tipologías requiere manejar información del paciente para el ajuste con el modelo y su posterior predicción.

MÉTODO

Dado que la investigación se basó exclusivamente en datos secundarios previamente recopilados y anonimizados, no se requirió la obtención directa del consentimiento informado por parte de los pacientes. Sin embargo, se garantizó en todo momento la confidencialidad, el resguardo de la identidad de los sujetos y el uso exclusivo de la información con fines científicos. La validación ética del protocolo asegura la integridad del estudio, considerando la naturaleza sensible del trastorno afectivo bipolar como objeto de análisis.

Métodos basados en algoritmos de inteligencia artificial

La elección de los modelos de clasificación empleados en este estudio sobre el trastorno afectivo bipolar en Ecuador responde a la necesidad de abordar un problema clínico complejo desde un enfoque computacional robusto, flexible y adaptado al contexto local. Se optó por algoritmos ampliamente validados como la regresión logística, Random Forest, SVM, árboles de decisión, K-vecinos más cercanos y Gradient Boosting debido a sus fortalezas específicas para el análisis de datos clínicos, que incluyen características como el manejo de muestras pequeñas, la presencia de datos faltantes, el ruido inherente a los registros médicos y la necesidad de interpretar modelos para su aplicación en entornos de salud mental.

La regresión logística ofrece una aproximación estadística tradicional útil por su interpretabilidad, mientras que Random Forest y Gradient Boosting destacan por su precisión y capacidad para evitar el sobreajuste. Las SVM fueron seleccionadas por su alto desempeño con pocos datos y su capacidad para identificar patrones no lineales, mientras que los árboles de decisión y KNN aportan transparencia y simplicidad en la toma de decisiones. Estos modelos permiten identificar factores de riesgo y clasificar estados clínicos del paciente, lo que resulta fundamental para mejorar la detección temprana y personalización del tratamiento en poblaciones vulnerables. Asimismo, se aplicó una validación cruzada rigurosa para comparar el rendimiento de los modelos y asegurar su capacidad de generalización, garantizando así que los resultados del análisis puedan ser utilizados con confianza en futuras aplicaciones clínicas y de investigación en el contexto ecuatoriano.

A continuación, se enumeran algunos de los algoritmos de aprendizaje de clasificación:

1) Regresión logística: representa una técnica estadística con enfoque para tareas de clasificación binaria que calcula la probabilidad de un resultado en función de los atributos de entrada. Este algoritmo de regresión logística estuvo sustentado con el uso de las bibliotecas Scikit-learn, TensorFlow y PyTorch.

2) Random Forest: representa una técnica de aprendizaje conjunto que reduce el sobreajuste y aumenta la precisión general mediante la construcción de varios árboles de decisión y la combinación de sus predicciones. La metódica Random Forest puede gestionar los datos faltantes y estos son resistentes al ruido y a los valores atípicos bajo librerías específicas TensorFlow y PyTorch.

3) Máquinas de vectores de soporte: representa una técnica potente que busca identificar el hiperplano ideal que divide dos clases para aplicaciones de clasificación y regresión. Las SVM son reconocidas por sus habilidades superiores de generalización y pueden manejar límites de decisión no lineales utilizando funciones del kernel. Se aplicó el algoritmo SVM como clasificador debido a su excelente rendimiento en un tamaño de muestra pequeño [15] mediante la adopción del paquete Scikit-learn basado en Python.[16]

4) Decision Tree: un árbol de decisión es un tipo de método de aprendizaje supervisado no paramétrico utilizado para tareas de clasificación y regresión. Se caracteriza por una estructura de árbol jerárquica que consta de un nodo raíz, ramas, nodos internos y nodos hoja. Python utiliza Scikit-learn, un paquete de aprendizaje automático de código abierto que incluye varios métodos de aprendizaje automático, como árboles de decisión.

5) K-Vecinos más cercanos: representa un enfoque no paramétrico para clasificar instancias utilizando el voto mayoritario de sus k vecinos más cercanos. KNN es fácil de entender y puede tener éxito con conjuntos de datos pequeños. Sin embargo, puede resultar computacionalmente costoso para grandes conjuntos de datos. En Python, se logra implementar este algoritmo usando la biblioteca scikit-learn.

6) Gradient Boosting: en el contexto de minimizar una función de pérdida, el aumento de gradiente representa un método de gradiente funcional que elige continuamente una función que apunta en la dirección de una hipótesis débil o un gradiente negativo. Se crea un modelo de predicción sólido combinando muchos modelos de aprendizaje débiles con un clasificador que aumenta el gradiente. Esta técnica de aprendizaje conjunto que construye árboles de decisión de forma secuencial y combina sus predicciones para mejorar el rendimiento general bajo algoritmo de Gradient Boosting a través de su clase GradientBoostingClassifier. Su excelente precisión predictiva lo ha convertido en una opción popular

en las implementaciones de aprendizaje automático.

Finalmente, se realizó una validación cruzada para producir un procedimiento de clasificación correcto, seguida de pruebas estadísticas para evaluar qué técnica de estimación funciona mejor. La variación cruzada antes mencionada compara la proporción de clasificación interna y externa correcta para evaluar la generalización de cada enfoque utilizado en la estimación del modelo.

Evaluación del modelo

Investigaciones relevantes en este campo revelan una amplia gama de métricas que pueden reflejar diversos sinónimos en los dominios de la medicina y ciencia de datos (recuperación de información). Incluimos aquí un breve resumen de las métricas más significativas basadas en el porcentaje de clasificación correcta, exactitud y precisión.[5] Las métricas en áreas de riesgo específicas incluyeron sensibilidad, especificidad y valor predictivo positivo.[5] Para facilitar las comparaciones indirectas con investigaciones anteriores, es aconsejable proporcionar tantas mediciones como sea posible para comparaciones futuras en el mismo campo de estudio.

RESULTADOS

Análisis exploratorios de datos preliminares

El análisis exploratorio de la variable prevalencia anual en el Ecuador de los trastornos afectivos bipolares, se encuentra descrito con una proporción aproximada del 0,94 % dentro de la cobertura temporal con una variabilidad en términos al parámetro prevalencia anual de 0,036. Esta métrica se obtuvo al manipular información de Our World in Data, la cual permite trabajar con valores estandarizados por edad para permitir comparaciones entre países y a lo largo del tiempo.[17,18]

Por el contrario, se puede trabajar a nivel porcentual en la evolución de la prevalencia general en trastornos afectivos bipolar en el Ecuador durante los últimos 22 años (figura 1). Cuyo comportamiento ha tenido una tendencia se ha incrementado en un pequeño diferencial. Sin embargo, durante el año 2020 hasta 2022 se produjo un incremento considerable en la prevalencia porcentual de casos en trastornos afectivos bipolar en el Ecuador.

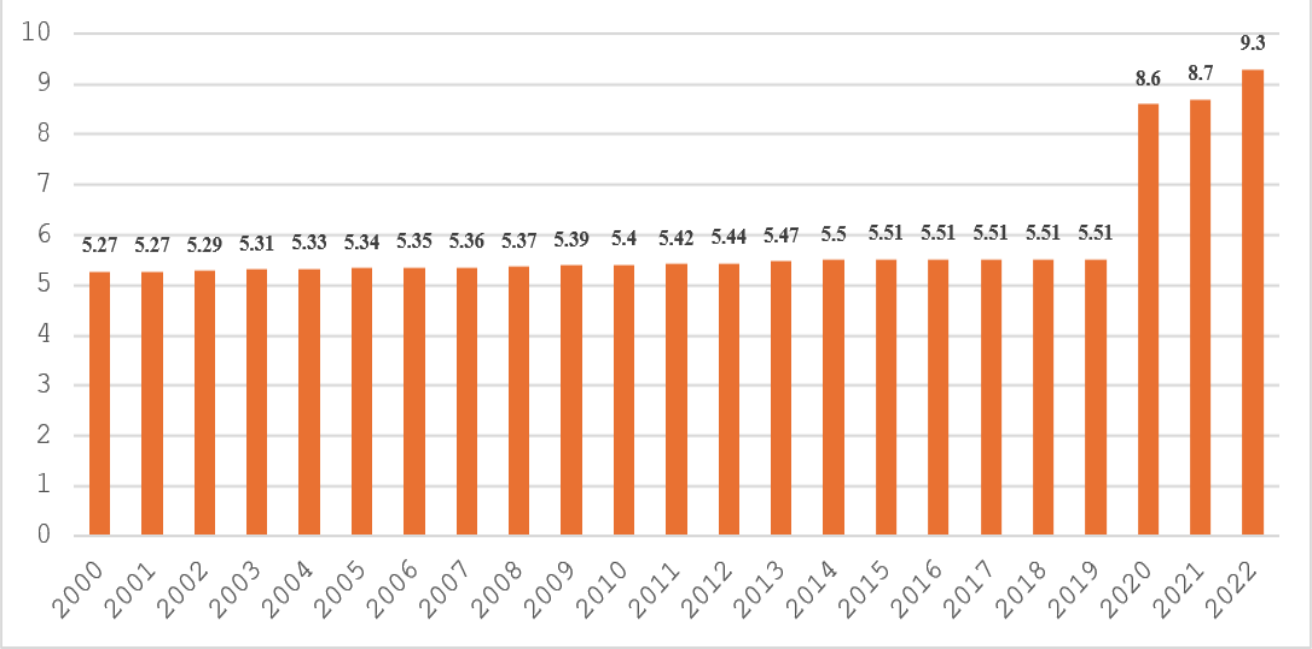

Figura 1. Prevalencia Porcentual de trastornos afectivos bipolar en Ecuador. Periodo 2000-2022
Fuente: Instituto Nacional de Estadística y Censos- Presidencia de la República del Ecuador.

Análisis Comparativo

En este apartado se desarrolló un análisis comparativo entre los modelos especificados mediante algoritmo de aprendizaje automático para clasificar y predecir trastorno afectivo bipolar en Ecuador (tabla 1). Estos modelos se construyeron con datos de corte transversal que corresponden a un total de 146 pacientes bajo monitoreo de los departamentos de psiquiatría adscrito al IESS.

Tabla 1. Representación de la exactitud comparativa en los modelos identificados

Modelo Estimado con validación cruzada	Exactitud en la clasificación
Logistic Regression (LR)	66,09 % (+/- 4,58)
Random Forest (RF)	89,35 % (+/- 3,85)
SVM	77,39 % (+/- 5,21)
Decision Tree (DT)	88,70 % (+/- 3,74)
K-Nearest Neighbors (KNN)	84,35 % (+/- 1,11)
Gradient Boosting (GB)	90,65 % (+/- 3,93)

Fuente: Python 3.11

Se visualiza en la tabla 1, lo concerniente al rendimiento en los modelos construidos considerando la métrica de precisión como medida que se calcula con la puntuación media del modelo, lo cual refleja el rendimiento del modelo. En este punto, se destacan dos posibles modelos que optimizan el rendimiento, Random Forest (89,35 %) y Gradient Boosting (90,65 %).

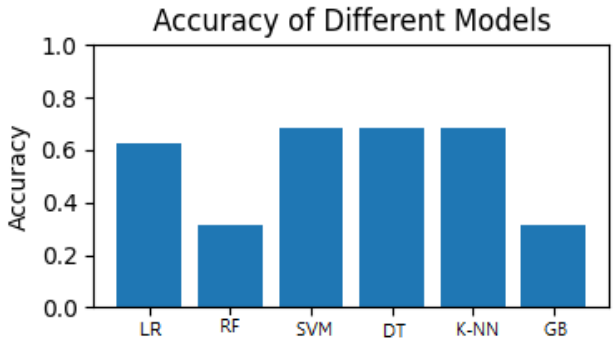

Figura 2. Representación de la precisión en los diferentes modelos

Según los resultados obtenidos, que describen que los modelos SVM, DT, K-NN y LR, los mismos presentan igual rendimiento al momento de discriminar pacientes con trastornos afectivos bipolar. En complemento, el modelo GB y RF representan la mejor opción para modelizar y explicar la presencia de este fenómeno (figura 2). Estos resultados se proporcionan mediante la implementación de los algoritmos en Python 3.11 bajo en el entorno de desarrolladores PyCharm 2024.1.

Además, se puede utilizar la desviación estándar de las puntuaciones multiplicada por 100 y formateada con dos decimales para comparar la variabilidad inherente al conjunto de datos y poder evaluar la precisión de un determinado modelo. En este sentido, entre los dos modelos identificados en la tabla 1 tienden a optimizar el rendimiento, se inclina por destacar los hallazgos en la elección de la estimación por Random Forest, los cuales presentan menor desviación estándar de las puntuaciones (+/- 3,85) en comparación con las obtenidas por la estimación con el modelo Gradient Boosting (+/- 3,93).

Con relación a la estimación de los mejores parámetros para el modelo Random Forest, se describen los mejores parámetros activando ('bootstrap': True) que representa un muestreo con reposición, lo que ayuda a mejorar la precisión del modelo. Luego, se limita a 10 la profundidad máxima en el modelo Random Forest, evitando el sobreajuste ('max_depth': 10), definiendo el número mínimo de muestras que deben estar presentes en una hoja para que el árbol se divida ('min_samples_leaf': 4); así como también, se estipula el número mínimo de cuatro muestras que deben estar presentes en un nodo para que el árbol se divida ('min_samples_split': 10) para obtener 10 estimadores como los mejores parámetros del modelo Random Forest (tabla 2).

Tabla 2. Importancia futura en la Clasificación con modelo Random Forest

Variable	Contribución
Age	0,432777
Marriage	0,207413
Days	0,130806
Work	0,074981

Fuente: Python 3.11

Los hallazgos dictaminan que la edad y el estado civil en la cohorte bajo estudio para clasificar el trastorno afectivo bipolar en el Ecuador presentan la mayor incidencia o peso para explicar la prevalencia de este tipo de enfermedades mediante el modelo Random Forest (tabla 2), el cual permite a través de los datos estimar parámetros de sensibilidad, precisión y puntuaciones de precisión. La sensibilidad definida por la métrica Recall, representar las frecuencias probabilísticas en torno a 0,89 para la detección del trastorno en conjunción con la verdadera tasa de positivos. Siendo esta la proporción de muestras que se identifican correctamente como positivas entre todas las muestras positivas existentes. La métrica de precisión, también conocida como frecuencia de valor predictivo positivo, indica la proporción de 0,81 en las muestras devueltas que discriminan una precisión positiva. La puntuación f1_score de 0,85, representa una métrica de precisión en la prueba que se logra por medio del cálculo en la media armónica entre la precisión y la recuperación. En función de lo anterior, el criterio de salida en términos a la construcción de la matriz de confusión (figura 3):

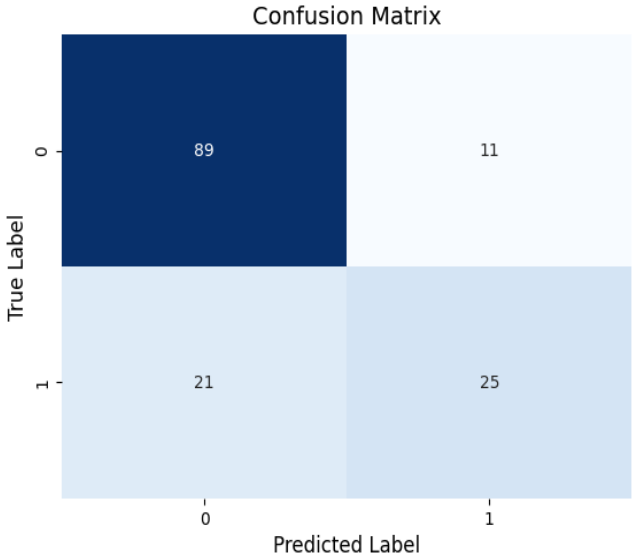

Figura 3. Matriz de confusión modelo Random Forest

Según los resultados que exhibe la matriz de confusión, el valor estimado en casos que logra una clasificación correcta en afecciones afectivas bipolar dentro de la cohorte bajo estudio se ubica en un 17 % (casos verdaderos positivos 25). A diferencia, se presenta el valor estimado de forma correcta de los casos verdaderos negativos o de no presentar trastornos afectivos bipolar 61 % (casos verdaderos negativos 89). En esencia, se tiene un 78 % de clasificación correcta, lo que se traduce en la representación del número de muestras correctamente identificadas bajo la población objeto de estudio. En este contexto, se discrimina las anteriores clasificaciones dispuestas en la figura 4, según las tipológicas afectivas bipolares definidas (Unipolar depresión, Bipolar I y Bipolar II).

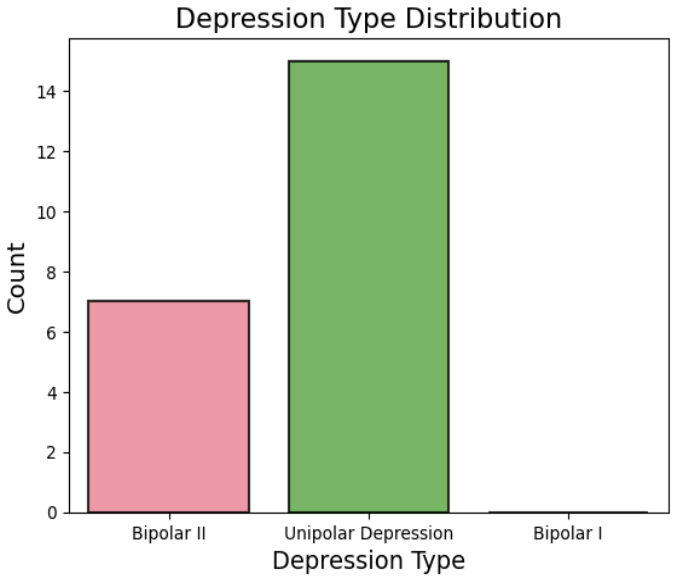

Figura 4. Clasificación de trastornos afectivo bipolar

Según lo que se muestra en la figura 4, se puede observar en términos al total en los 146 pacientes sometidos a estudio, 15 de ellos entran en la clasificación de un cuadro en trastornos unipolar depresivos que refleja una incidencia de la enfermedad en un 10 % según la población muestreada, en adición se presentan siete (07) individuos que se acopla a un cuadro de trastorno bipolar tipo II; lo que equivale a una prevalencia del 5 % y sólo tres (03) individuos que se definen entrar en un cuadro bipolar tipo I (prevalencia en un 2 %).

DISCUSIÓN

El presente estudio, enfocado en el modelado matemático del trastorno afectivo bipolar en Ecuador, presenta diversas limitaciones que deben ser consideradas al interpretar sus resultados. En primer lugar, la disponibilidad y calidad de los datos clínicos representa un desafío, ya que en Ecuador no existe una base de datos nacional estandarizada que registre de forma sistemática los casos de trastornos afectivos. Esto obliga a utilizar fuentes secundarias o datos institucionales con posible subregistro o sesgos de acceso, especialmente en zonas rurales o de difícil acceso. Además, el modelo matemático requiere simplificar dinámicas complejas asociadas al trastorno bipolar, como comorbilidades psiquiátricas, impacto del entorno familiar, adherencia terapéutica o exposición a eventos estresantes, los cuales no pueden modelarse completamente debido a su carácter multifactorial y subjetivo. Asimismo, los parámetros epidemiológicos utilizados se mantienen estáticos en el tiempo, sin considerar fluctuaciones sociales, económicas o sanitarias que podrían alterar significativamente la incidencia, el curso clínico o la respuesta terapéutica de la población. Otra limitación importante es la representación geográfica parcial del modelo, dado que la mayoría de los datos provienen de zonas urbanas con mayor acceso a servicios de salud mental, lo cual excluye realidades de poblaciones indígenas o comunidades marginadas.

A pesar de estar enfocado en la población ecuatoriana, el modelo matemático desarrollado tiene un fuerte potencial de replicabilidad en otros contextos latinoamericanos, especialmente en países que comparten condiciones estructurales similares en salud mental. El marco de modelado utilizado, basado en sistemas dinámicos y tasas de transición entre estados clínicos, puede ser fácilmente adaptado a distintas poblaciones mediante la actualización de parámetros epidemiológicos específicos, como prevalencia, duración de episodios, tasas de recaída o adherencia terapéutica. Esta característica lo convierte en una herramienta útil no solo para análisis locales, sino también para estudios comparativos regionales.

El modelo Random Forest exhibe una alta sensibilidad, lo que indica un alto grado de capacidad para la detección de casos significativos. Esto indica que tiene excelentes habilidades para discriminar casos negativos. En otras palabras, la generación de falsos positivos es poco probable. Como resultado, este modelo es mucho más sensible que específico. Cuando nuestro objetivo es evitar los falsos positivos a toda costa, este es el escenario que llama nuestra atención.

El modelo Random Forest (RF) representa una técnica de selección de características de tipo integrado que se utiliza a menudo en la minería de datos,[19] se ha demostrado que aumenta considerablemente el rendimiento del modelo,[20,21] se identifica genes clave.[22] Una forma de medir la importancia de una característica en el modelo de RF es utilizar la decadencia media de impurezas (MDI). Es la reducción media de la impureza de los nodos en todos los árboles, ponderada por la probabilidad de llegar a ese nodo. En este trabajo, utilizamos una metodología de aprendizaje supervisado para construir un modelo de bosque aleatorio en el conjunto de entrenamiento utilizando la configuración predeterminada.

En concordancia, Llamocca et al.[23] demostraron un afecto negativo mejorado amplifica los contrastes emocionales positivos cuando se enfrentan a eventos positivos inesperados, mientras que las personas sensibles a los cambios emocionales negativos utilizan aumentos previos del afecto negativo para evitar una mayor escalada en respuesta a cambios adversos, según un modelo de evitación de contrastes que se desarrolló. Las personas que padecen trastornos del espectro bipolar, que se caracterizan por oscilaciones entre episodios de hipomanía y depresión, donde se pueden notar diferencias más sutiles en sus sentimientos.

Se identificó un impacto marginal en el subtipo bipolar tipo II y mostró una mayor propensión a evitar contrastes emocionales desagradables en comparación con la enfermedad bipolar tipo I según el modelo de evitación de contrastes. Estos resultados implican que un proceso psicológico relacionado con las enfermedades del espectro bipolar puede ser la evitación del contraste cuya detección temprana de personas en riesgo de padecer enfermedad afectiva bipolar permite un examen y una centrada prevención. Aunque se han identificado una variedad de factores de riesgo de enfermedad bipolar, como antecedentes en grupo de familiares [24] y acontecimientos vitales estresantes,[25] donde la predicción de riesgos de forma cuantitativa y escalable sigue siendo un desafío. Por ende, las investigaciones anteriores se han concentrado principalmente en aquellos que han tenido depresión en el pasado y/o han empleado pequeñas muestras.

Los estudios anteriores se han centrado en gran medida en individuos con antecedentes de depresión y/o han incluido muestras relativamente pequeñas.[26] En este desarrollo se partió considerando evaluar un contexto general de prevalencia de la enfermedad afectiva bipolar dentro de la República del Ecuador durante la cobertura temporal 2000-2022. En un contexto específico, se manejó una cohorte de personas con la

característica de interés en el presente estudio para lograr validar múltiples enfoques y describir el fenómeno bajo un diagnóstico de enfermedad afectiva bipolar.

En la práctica, los avances en la medicina clínica y de precisión podrían conducir al desarrollo de modelos de predicción de la enfermedad bipolar generalizables que hayan sido entrenados y validados en varios sistemas de salud. Si bien la precisión de los modelos es fuerte, no tiene en cuenta los factores sociales y ambientales innatos que influyen en el comportamiento afectivo bipolar. Más bien, describe y predice los componentes físicos y psicológicos de la calidad de vida.

Orsolini usó el fenotipado digital para clasificar el trastorno bipolar y el trastorno unipolar: hallazgos exploratorios mediante modelos de aprendizaje automático bajo los siguientes objetivos fueron investigar 1) disparidades entre el trastorno bipolar y el trastorno unipolar en los datos basados en teléfonos inteligentes sobre el uso del teléfono; y 2) categorización inteligente del trastorno bipolar y el trastorno unipolar mediante modelos de aprendizaje automático, sensibilidad, especificidad y AUC de los datos telefónicos. Durante seis meses, hubo acceso a autoevaluaciones diarias del estado de ánimo basadas en teléfonos inteligentes y, al mismo tiempo, se recopilaron pasivamente datos sobre el uso de teléfonos inteligentes. En resumen, el fenotipado digital subraya la dificultad de generalizar a personas con pocos elementos tangibles, incluso si resulta prometedor a la hora de distinguir entre pacientes con enfermedad bipolar y enfermedad unipolar. Debe utilizarse junto con evaluaciones clínicas exhaustivas de los profesionales.[27]

El desarrollo de modelos de predicción generalizables del riesgo de enfermedad bipolar en una variedad de circunstancias empíricas es, en última instancia, el camino hacia la medicina de precisión. Una estrategia de conjunto demostró ser la más efectiva en general cuando se compararon una variedad de algoritmos de aprendizaje automático; no obstante, requirió un reciclaje local, en ese sentido se utilizaron diferentes algoritmos (bosques aleatorios, máquinas de aumento de gradiente, regresión penalizada y algoritmos de aprendizaje de conjuntos apilados) para crear modelos de predicción válidos.

En este contexto, un estudio reciente abordó predictores que se limitaron a atributos derivados de registros médicos electrónicos de acceso público, sin referencia a un modelo de datos compartido que abarcara datos demográficos, códigos de enfermedades y medicamentos recetados, que demuestra que la detección temprana del riesgo de trastorno bipolar utilizando modelos de predicción generalizables entrenados en diversas cohortes en poblaciones ecuatorianas debería mejorar la evaluación enfocada de los pacientes de alto riesgo, prevenir diagnósticos erróneos y optimizar la asignación de recursos escasos de salud mental.[28]

CONCLUSIONES

El análisis exploratorio de la variable prevalencia anual en el Ecuador de los trastornos afectivos bipolares se encuentra descrito con una proporción aproximada del 0,94 % dentro de la cobertura temporal. Esta métrica se obtuvo al manipular información de Our World in Data, que permite trabajar con valores estandarizados por edad para permitir comparaciones entre países y a lo largo del tiempo. En cuanto al análisis comparativo entre los modelos especificados mediante algoritmo de aprendizaje automático para clasificar y predecir trastorno afectivo bipolar en Ecuador. Los modelos SVM, DT, K-NN y LR presentan igual rendimiento al momento de discriminar pacientes con trastornos afectivos bipolar. En complemento, el modelo GB y RF representan la mejor opción para modelizar y explicar la presencia de este fenómeno.

La importancia futura en la clasificación con modelo Random Forest se enfoca en la mayor incidencia o peso para explicar la prevalencia de este tipo de enfermedades. Este modelo Random Forest permite estimar parámetros de sensibilidad, precisión y puntuaciones de precisión. La sensibilidad definida por la métrica Recall representa las frecuencias probabilísticas en torno a 0,89 para la detección del trastorno en conjunción con la verdadera tasa de positivos. La precisión, también conocida como frecuencia de valor predictivo positivo, indica la proporción de 0,81 en las muestras devueltas que discriminan una precisión positiva. La puntuación f1_score de 0,85 representa una métrica de precisión en la prueba que se logra por medio del cálculo en la media armónica entre la precisión y la recuperación. O, efectivamente, considerar a ampliar el estudio evaluando mediante inteligencia artificial los registros actigráficos de la actividad motora, los cuales representan una forma objetiva de observar la depresión, incluso si las investigaciones psiquiátricas sobre este tema no lo han analizado con gran detalle.

REFERENCIAS BIBLIOGRÁFICAS

1. OMS. Psychosis and bipolar disorders. Who.int. Disponible en: https://www.who.int/teams/mental-health-and-substance-use/treatment-care/mental-health-gap-action-programme/evidence-centre/psychosis-

2. Weye N, Momen NC, Christensen MK, Iburg KM, Dalsgaard S, Laursen TM, et al. Association of specific mental disorders with premature mortality in the danish population using alternative measurement methods. JAMA Netw Open. 2020;3(6):e206646. http://dx.doi.org/10.1001/jamanetworkopen.2020.6646

3. Biazus TB, Beraldi GH, Tokeshi L, Rotenberg L de S, Dragioti E, Carvalho AF, et al. All-cause and cause-specific mortality among people with bipolar disorder: a large-scale systematic review and meta-analysis. Mol Psychiatry. 2023;28(6):2508-24. http://dx.doi.org/10.1038/s41380-023-02109-9

4. McIntyre RS, Alda M, Baldessarini RJ, Bauer M, Berk M, Correll CU, et al. The clinical characterization of the adult patient with bipolar disorder aimed at personalization of management. World Psychiatry. 2022;21(3):364-87. http://dx.doi.org/10.1002/wps.20997

5. Chakrabarti S, Singh N. Psychotic symptoms in bipolar disorder and their impact on the illness: A systematic review. World J Psychiatry. 2022;12(9):1204-32. http://dx.doi.org/10.5498/wjp.v12.i9.1204

6. Holm M, Taipale H, Tanskanen A, Tiihonen J, Mitterdorfer-Rutz E. Employment among people with schizophrenia or bipolar disorder: A population-based study using nationwide registers. Acta Psychiatr Scand. 2021;143(1):61-71. http://dx.doi.org/10.1111/acps.13254

7. Joas E, Bäckman K, Karanti A, Sparding T, Colom F, Pålsson E, et al. Psychoeducation for bipolar disorder and risk of recurrence and hospitalization - a within-individual analysis using registry data. Psychol Med. 2020;50(6):1043-9. http://dx.doi.org/10.1017/s0033291719001053

8. Walsh CG, Ripperger MA, Hu Y, Sheu Y-H, Lee H, Wilimitis D, et al. Development and multi-site external validation of a generalizable risk prediction model for bipolar disorder. Transl Psychiatry. 2024;14(1). http://dx.doi.org/10.1038/s41398-023-02720-y

9. Rantala MJ, Luoto S, Borráz-León JI, Krams I. Bipolar disorder: An evolutionary psychoneuroimmunological approach. Neurosci Biobehav Rev. 2021;122:28-37. http://dx.doi.org/10.1016/j.neubiorev.2020.12.031

10. Cheon E-J, Bearden CE, Sun D, Ching CRK, Andreassen OA, Schmaal L, et al. Cross disorder comparisons of brain structure in schizophrenia, bipolar disorder, major depressive disorder, and 22q11.2 deletion syndrome: A review of ENIGMA findings. Psychiatry Clin Neurosci. 2022;76(5):140-61. http://dx.doi.org/10.1111/pcn.13337

11. Malhi GS, Bell E, Bassett D, Boyce P, Bryant R, Hazell P, et al. The 2020 Royal Australian and New Zealand College of Psychiatrists clinical practice guidelines for mood disorders. Aust N Z J Psychiatry. 2021;55(1):7-117. http://dx.doi.org/10.1177/0004867420979353

12. Di Salvo G, Porceddu G, Albert U, Maina G, Rosso G. Correlates of long duration of untreated illness (DUI) in patients with bipolar disorder: results of an observational study. Ann Gen Psychiatry. 2023;22(1). http://dx.doi.org/10.1186/s12991-023-00442-5

13. IHME Bipolar disorder prevalence. Our World in Data. Disponible en: https://ourworldindata.org/grapher/bipolar-disorder-prevalence?tab=chart&time=earliest..latest&country=~ECU

14. INEC Ecuador - Estadísticas Hospitalarias Camas y Egresos 2022. Gob.ec. Disponible en: https://anda.inec.gob.ec/anda/index.php/catalog/976/datafile/F47/V1715

15. Carruthers SP, Rossell SL, Murray G, Karantonis J, Furlong LS, Van Rheenen TE. Mindfulness, mood symptom tendencies and quality of life in bipolar disorder: An examination of the mediating influence of emotion regulation difficulties. J Affect Disord. 2022;298:166-72. http://dx.doi.org/10.1016/j.jad.2021.10.107

16. Lalli M, Brouillette K, Kapczinski F, de Azevedo Cardoso T. Substance use as a risk factor for bipolar disorder: A systematic review. J Psychiatr Res. 2021;144:285-95. http://dx.doi.org/10.1016/j.jpsychires.2021.10.012

17. Rodriguez V, Alameda L, Trotta G, Spinazzola E, Marino P, Matheson SL, et al. Environmental risk factors in Bipolar disorder and Psychotic depression: A systematic review and meta-analysis of prospective studies. Schizophr Bull. 2021;47(4):959-74. http://dx.doi.org/10.1093/schbul/sbaa197

18. He H, Hu C, Ren Z, Bai L, Gao F, Lyu J. Trends in the incidence and DALYs of bipolar disorder at global, regional, and national levels: Results from the global burden of Disease Study 2017. J Psychiatr Res. 2020;125:96-105. http://dx.doi.org/10.1016/j.jpsychires.2020.03.015

19. Palacios-Ariza MA, Morales-Mendoza E, Murcia J, Arias-Duarte R, Lara-Castellanos G, Cely-Jiménez A, et al. Prediction of patient admission and readmission in adults from a Colombian cohort with bipolar disorder using artificial intelligence. Front Psychiatry. 2023;14. http://dx.doi.org/10.3389/fpsyt.2023.1266548

20. Montazeri M, Montazeri M, Bahaadinbeigy K, Montazeri M, Afraz A. Application of machine learning methods in predicting schizophrenia and bipolar disorders: A systematic review. Health Sci Rep. 2023;6(1). http://dx.doi.org/10.1002/hsr2.962

21. Wanderley Espinola C, Gomes JC, Mônica Silva Pereira J, dos Santos WP. Detection of major depressive disorder, bipolar disorder, schizophrenia and generalized anxiety disorder using vocal acoustic analysis and machine learning: an exploratory study. Res Biomed Eng. 2022;38(3):813-29. http://dx.doi.org/10.1007/s42600-022-00222-2

22. Cearns M, Amare AT, Schubert KO, Thalamuthu A, Frank J, Streit F, et al. Using polygenic scores and clinical data for bipolar disorder patient stratification and lithium response prediction: machine learning approach. Br J Psychiatry. 2022;220(4):219-28. http://dx.doi.org/10.1192/bjp.2022.28

23. Llamocca P, López V, Santos M, Čukić M. Personalized characterization of emotional states in patients with bipolar disorder. Mathematics. 2021;9(11):1174. http://dx.doi.org/10.3390/math9111174

24. Gordovez FJA, McMahon FJ. The genetics of bipolar disorder. Mol Psychiatry. 2020;25(3):544-59. http://dx.doi.org/10.1038/s41380-019-0634-7

25. Mikolas P, Bröckel K, Vogelbacher C, Müller DK, Marxen M, Berndt C, et al. Individuals at increased risk for development of bipolar disorder display structural alterations similar to people with manifest disease. Transl Psychiatry. 2021;11(1). http://dx.doi.org/10.1038/s41398-021-01598-y

26. Comsa M, Anderson KN, Sharma A, Yadav VC, Watson S. The relationship between sleep and depression and bipolar disorder in children and young people. BJPsych Open. 2022;8(1). http://dx.doi.org/10.1192/bjo.2021.1076

27. Orsolini L, Fiorani M, Volpe U. Digital phenotyping in bipolar disorder: Which integration with clinical endophenotypes and biomarkers? Int J Mol Sci. 2020;21(20):7684. http://dx.doi.org/10.3390/ijms21207684

28. Sonkurt HO, Altınöz AE, Çimen E, Köşger F, Öztürk G. The role of cognitive functions in the diagnosis of bipolar disorder: A machine learning model. Int J Med Inform. 2021;145(104311):104311. http://dx.doi.org/10.1016/j.ijmedinf.2020.104311

FINANCIACIÓN

Los autores no recibieron financiación para el desarrollo de la presente investigación.

CONFLICTO DE INTERESES

Los autores declaran que no existe conflicto de intereses.

CONTRIBUCIÓN DE AUTORÍA

Conceptualización: Cristhian Ismael Gómez Gaona.
Curación de datos: Andrea del Rocío Mejía Rubio.
Análisis formal: José Rubén León Pérez.
Investigación: Jesús Rodríguez.
Metodología: Zilma Diago Alfes.
Administración del proyecto: Laura Esther Muñoz Escobar.
Recursos: Cristian Inca.
Software: Jimmy Yaguana Torres.
Supervisión: Marco Velasco Arellano.
Validación: Jesús Rodríguez.
Visualización: Cristian Gómez.
Redacción – borrador original: Cristian inca.
Redacción – revisión y edición: Marco Velasco Arellano.