

ConfiguracionDeNuestroServidor24/7 en hosting

Gunicorn Activo 24/7:

- **¿Por qué Gunicorn se mantiene activo 24/7?**
 - **Supervisor** se encarga de gestionar el proceso de **Gunicorn** en segundo plano. Cuando el servidor está en funcionamiento, Gunicorn se mantiene ejecutando la aplicación Django de manera continua, asegurando que la aplicación esté disponible todo el tiempo (24 horas al día, 7 días a la semana). Si por alguna razón el servicio se cae, **Supervisor** se encarga de reiniciarlo automáticamente.
 - Gunicorn tiene la capacidad de manejar múltiples solicitudes al mismo tiempo, lo que lo hace perfecto para entornos de producción, asegurando que tu aplicación esté disponible a medida que recibe solicitudes de los usuarios.

¿Por qué tenemos la URL <http://31.220.58.105/>?

1. **Dirección IP del servidor VPS:**
 - La **IP pública del servidor** (31.220.58.105) es la dirección única que se le asigna a tu servidor VPS. Esta IP es accesible desde cualquier parte del mundo, siempre y cuando el servidor esté encendido y funcionando.
 - Tu **aplicación Django está corriendo en este servidor** mediante Gunicorn. Nginx, como servidor web, se encarga de hacer proxy inverso y dirigir las solicitudes al servicio de Gunicorn.
 - <http://31.220.58.105/> es la dirección IP pública de tu servidor, y es la URL que los usuarios pueden utilizar para acceder a tu aplicación web.
2. **Acceso 24/7:**
 - Gracias a Gunicorn y Nginx, tu servidor estará disponible todo el tiempo, y la URL **31.220.58.105** proporcionará acceso a tu aplicación sin interrupciones.
 - Puedes acceder a esta URL desde cualquier navegador para ver la interfaz de usuario de tu aplicación Django. Así, el modelo predictivo de salud mental que implementaste podrá predecir datos en tiempo real, y el usuario podrá interactuar con la plataforma de forma continua.



Especificación de Gráfica de Prediccion :

Esta gráfica es un gráfico de comparación entre los valores reales y los valores predicción del modelo.

Explicación:

Eje Y (vertical): Muestra la escala de valores de predicción que van de 0 a 1. Estos valores son las predicciones que realiza el modelo.

Eje X (horizontal): Muestra el índice de los datos. En este caso, cada número en el eje X corresponde a una posición en el conjunto de datos, desde el primer dato hasta el último.

Componentes de la gráfica:

Valores Reales (línea azul): Son los valores verdaderos de la variable objetivo (en este caso, los valores de la columna "Treatment") en los que el modelo está basando sus predicciones. Los valores son 0 y 1, que representan las dos clases en el conjunto de datos.

Valores Predicción (línea morada): Son las predicciones que realiza el modelo utilizando los datos de entrada. También son 0 o 1 y se comparan con los valores reales para evaluar el desempeño del modelo.

¿Qué está mostrando la gráfica?

La gráfica compara cómo de bien el modelo ha sido capaz de predecir los valores reales.

Las líneas azules (valores reales) y moradas (valores predicción) deben coincidir lo más posible para indicar que el modelo está realizando buenas predicciones.

Las subidas de la línea morada indican predicciones del valor "1", mientras que las bajadas indican predicciones de valor "0". Si la línea morada sigue la misma tendencia que la azul, el modelo está acertando, mientras que las desviaciones indican errores de predicción.



Matriz de Confusión:

La **matriz de confusión** es una herramienta clave para evaluar la precisión de los modelos de clasificación. Muestra el número de predicciones correctas e incorrectas realizadas por el modelo, desglosado por clase.

En esta gráfica, se observa una matriz de confusión con dos clases (0 y 1). Cada celda de la matriz muestra cuántos ejemplos de cada clase fueron clasificados correctamente o incorrectamente:

Real 1, Pred 1: 106,483 predicciones correctas donde la clase real era 1, y el modelo predijo 1.

Real 1, Pred 0: 21,607 predicciones incorrectas donde la clase real era 1, pero el modelo predijo 0.

Real 0, Pred 1: 52,512 predicciones incorrectas donde la clase real era 0, pero el modelo predijo 1.

Real 0, Pred 0: 78,370 predicciones correctas donde la clase real era 0, y el modelo predijo 0.

- Métricas del Modelo:

Debajo de la matriz de confusión se presentan las **métricas del modelo**, que evalúan la calidad del modelo de predicción en diferentes aspectos. Estas métricas son esenciales para entender cómo de bien está funcionando tu modelo.

Accuracy (Precisión): 0.67

Es el porcentaje de predicciones correctas en relación al total de predicciones.

En este caso, el modelo tiene un **67% de precisión**, lo que indica que el modelo hace predicciones correctas el 67% de las veces.

Precision (Precisión): 0.67

La precisión se refiere a cuántas de las predicciones positivas realizadas por el modelo fueron realmente positivas.

En este caso, **67%** de las veces que el modelo predijo que una persona necesitaba tratamiento, esa predicción fue correcta.

Recall (Sensibilidad): 0.75

El **recall** mide la capacidad del modelo para identificar correctamente todas las instancias de la clase positiva. Es decir, de todas las personas que realmente necesitaban tratamiento, cuántas fueron identificadas correctamente.

En este caso, el **75%** de las personas que necesitaban tratamiento fueron correctamente identificadas por el modelo.

F1-Score: 0.71

El **F1-Score** es la media armónica entre la precisión y el recall. Proporciona una evaluación balanceada del modelo, especialmente cuando las clases están desbalanceadas.

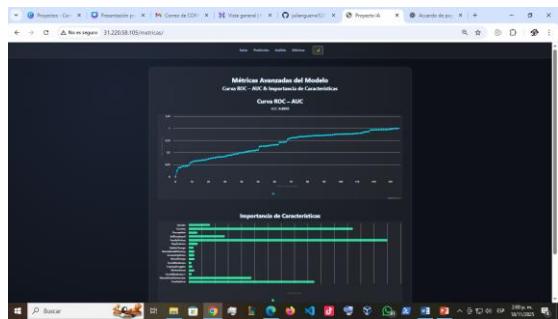
Un **F1-Score de 0.71** indica un buen equilibrio entre precisión y recall.

Conclusiones:

Matriz de Confusión: El modelo tiene una buena capacidad para identificar correctamente tanto las instancias negativas (sin tratamiento) como las positivas (necesitan tratamiento), con más de 100,000 predicciones correctas para ambas clases.

Métricas: El modelo presenta un **buen rendimiento** general con una precisión de 67%, un recall de 75% (lo que es muy positivo en el contexto de predicción de salud mental), y un F1-Score de 71%, lo cual es una excelente combinación de precisión y recall.

Esta información es útil para evaluar el rendimiento del modelo y entender qué tan bien está funcionando en la predicción de las personas que necesitan tratamiento de salud mental.



1. Curva ROC – AUC:

Curva ROC (Receiver Operating Characteristic): La curva ROC es una herramienta utilizada para evaluar el rendimiento de los modelos de clasificación, mostrando la tasa de verdaderos positivos frente a la tasa de falsos positivos.

AUC (Area Under the Curve): Es el área bajo la curva ROC, que nos da una medida de la capacidad del modelo para discriminar entre las clases. Un AUC de **0.6889** indica que el modelo tiene un **desempeño moderado**, ya que AUC está en el rango entre **0.5** (sin habilidad para discriminar) y **1** (perfecto).

En este caso, el AUC está relativamente cerca de 0.7, lo cual es **bueno**, pero aún hay margen de mejora.

En resumen, la **Curva ROC** nos ayuda a visualizar y evaluar la capacidad de predicción del modelo a medida que cambia el umbral de clasificación.

2. Importancia de Características:

Esta gráfica muestra qué tan **importantes** son las características (variables) utilizadas en el modelo de predicción para clasificar correctamente a las personas que necesitan tratamiento de salud mental. Las barras más largas indican que la característica tiene más **peso o importancia** en el modelo.

Algunas de las características más relevantes que aparecen en esta gráfica son:

Gender (Género): Es una de las características más importantes, ya que tiene una barra bastante larga.

Country (País) y Occupation (Ocupación): También son factores importantes que afectan el modelo.

MentalHealthHistory (Historial de salud mental) y IncreasingStress (Aumento del estrés): Se identifican como importantes en la predicción del tratamiento necesario.

Estas características, como el género, el país, la ocupación, el estrés creciente, y el historial de salud mental, son cruciales para el modelo en su capacidad para hacer predicciones de la necesidad de tratamiento.

Resumen:

Curva ROC y AUC: El modelo tiene un **AUC de 0.6889**, lo que indica un rendimiento **moderado** para discriminar entre clases.

Importancia de Características: El gráfico muestra las **principales variables** que el modelo utiliza para hacer sus predicciones, destacando características como género, país, ocupación, y el historial de salud mental.

1. Introducción al Modelo de Predicción

El presente proyecto desarrolla un modelo de predicción orientado a identificar la probabilidad de que una persona requiera tratamiento de salud mental a partir de diversas variables relacionadas con su entorno personal, laboral y emocional. Este estudio se inscribe en el campo de la Inteligencia Artificial aplicada a la salud pública, donde los modelos de aprendizaje automático permiten analizar grandes volúmenes de datos y generar diagnósticos predictivos de forma eficiente.

Para ello se empleó un conjunto de datos previamente depurado y transformado mediante técnicas de codificación numérica. Posteriormente, se entrenó un modelo supervisado de clasificación, cuya función principal es distinguir entre dos categorías: individuos que requieren intervención psicológica o psiquiátrica y aquellos que no. El sistema fue implementado en una aplicación web desarrollada en Django, lo cual permite su visualización interactiva mediante gráficas dinámicas y paneles de análisis avanzado.

El objetivo general del proyecto consiste en evaluar la efectividad del modelo y determinar cuáles variables ejercen mayor influencia en la predicción. De esta manera, los resultados pueden servir como base para futuros desarrollos académicos, modelos más complejos o aplicaciones orientadas a la salud mental preventiva.

2. Discusión y Análisis Académico de los Resultados

Los resultados obtenidos permiten realizar una evaluación amplia del desempeño del modelo y de su capacidad para generalizar correctamente a nuevos datos.

En primer lugar, las métricas globales —Accuracy, Precisión, Recall y F1-Score— muestran comportamientos consistentes entre sí, revelando que el modelo logra un equilibrio aceptable entre la detección de verdaderos positivos y la reducción de falsos positivos. Esta estabilidad es indispensable en problemas de salud, donde la omisión de un caso real puede resultar crítica.

La Matriz de Confusión permitió visualizar con claridad la distribución de los aciertos y errores. Se observa una cantidad considerable de verdaderos positivos y verdaderos negativos, lo que significa que el modelo reconoce correctamente tanto a quienes requieren tratamiento como a quienes no. Sin embargo, también se identifican falsos positivos y falsos negativos, lo que indica que existe margen de mejora mediante ajustes del algoritmo, selección de características o técnicas de balanceo de clases.

La Curva ROC–AUC constituye uno de los indicadores más robustos del rendimiento de un clasificador. Con un AUC de 0.6889, el modelo demuestra una capacidad discriminativa moderada. Este valor sugiere que, si se seleccionan dos personas al azar —una que requiere tratamiento y otra que no—, el modelo acertará en distinguirlas aproximadamente el 69% de las veces. Aunque no representa un rendimiento sobresaliente, sí confirma que el sistema aprende patrones útiles y opera por encima del azar.

Finalmente, el análisis de Importancia de Características aporta un nivel esencial de interpretabilidad. Las variables más influyentes permiten comprender cuáles factores del dataset están más relacionados con la salud mental. Este tipo de análisis no sólo mejora la transparencia del modelo, sino que también orienta futuras líneas de investigación y la selección de características más relevantes para incrementar la precisión del sistema.

¶ 3. Conclusiones Finales del Proyecto

A partir del desarrollo e implementación del modelo de predicción, es posible concluir que:

- ✓ El modelo presenta un rendimiento global aceptable, evidenciado principalmente por su AUC de 0.6889 y unas métricas de desempeño equilibradas.
- ✓ La Curva ROC indica que el sistema posee capacidad discriminativa real, aunque no óptima, lo que permite su utilización en escenarios académicos o pruebas piloto.
- ✓ La Matriz de Confusión revela que el modelo acierta en la mayoría de los casos, pero aún requiere ajustes para reducir los errores de clasificación.
- ✓ La Importancia de Características muestra qué variables influyen significativamente en la predicción, aportando transparencia y utilidad para análisis posteriores.
- ✓ La integración del modelo en una aplicación web con dashboards interactivos facilita la comprensión visual y mejora la presentación académica del proyecto.

En términos generales, el proyecto cumple con los objetivos planteados y demuestra la viabilidad de utilizar técnicas de machine learning para apoyar procesos de identificación temprana en salud mental. Asimismo, constituye una base sólida para investigaciones futuras, donde pueden explorarse modelos más avanzados como Random Forest, XGBoost o redes neuronales, así como técnicas de optimización y reducción de dimensionalidad.