

# Inteligencia Artificial II

Juan Pablo Restrepo Uribe

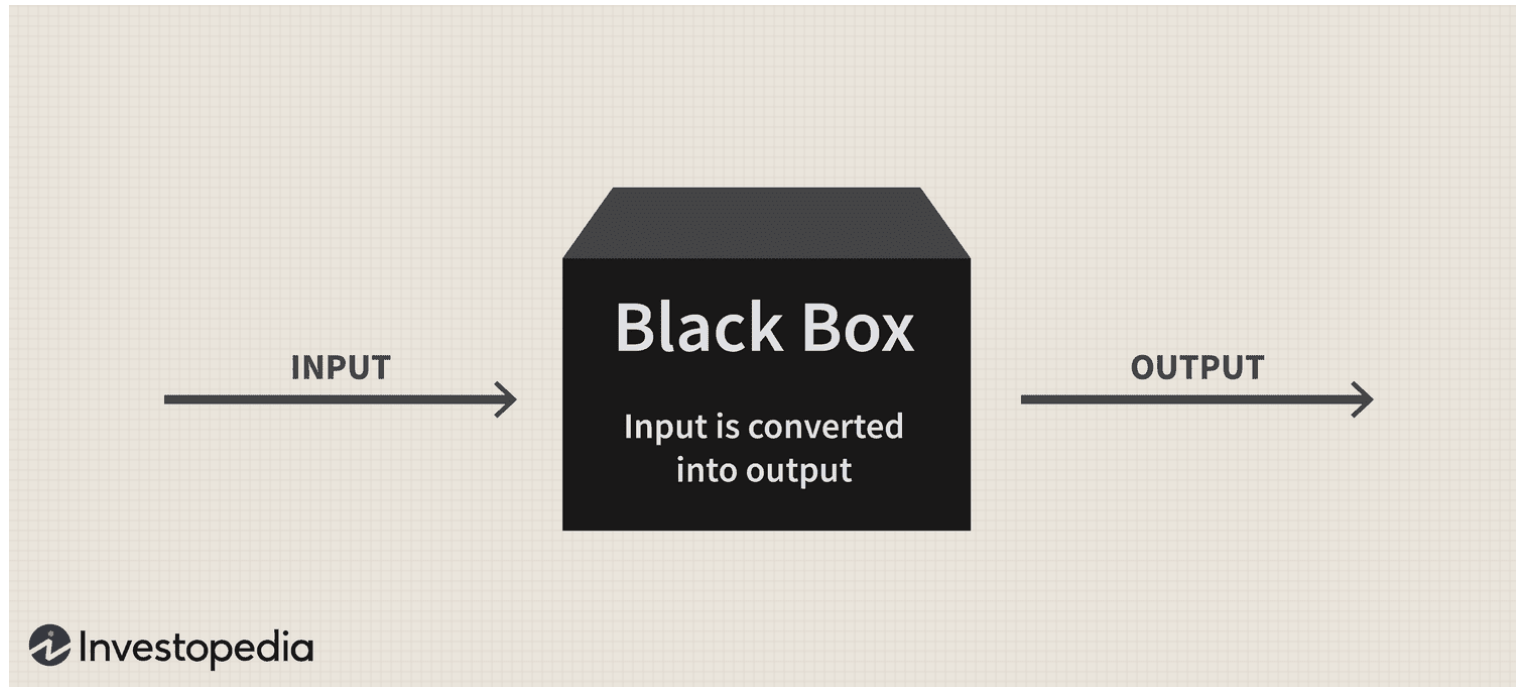
Ing. Biomedico - MSc. Automatización y Control Industrial

[jprestrepo@correo.iue.edu.co](mailto:jprestrepo@correo.iue.edu.co)

2023

Institución Universitaria de Envigado

## Caja negra



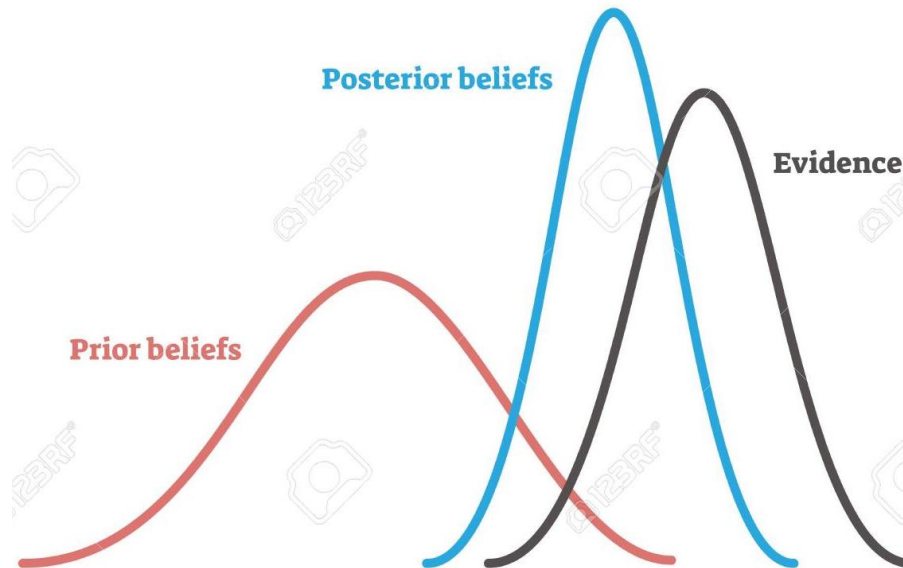
En términos prácticos, el problema de la caja negra de la IA es la dificultad de descifrar el razonamiento que subyace a las predicciones o decisiones de un sistema de IA.

## Caja negra

"En los últimos años, se ha prestado mucha atención al desarrollo de técnicas para interpretar y explicar las decisiones tomadas por los modelos de IA, como la generación de puntuaciones de importancia de las características, la visualización de los límites de las decisiones y la identificación de explicaciones hipotéticas contra fácticas"

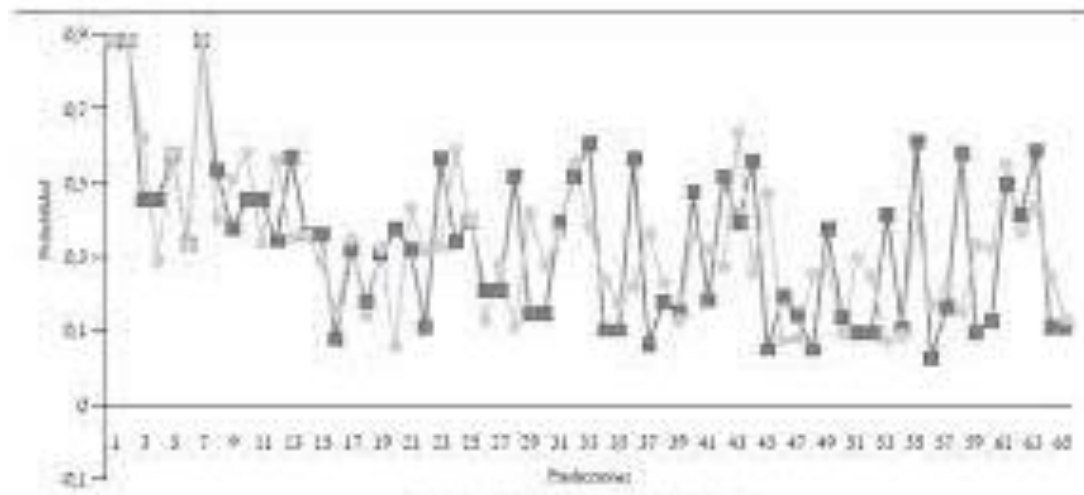
# Optimización bayesiana

“La Optimización Bayesiana (OB) es un campo con una creciente popularidad actualmente debido a su utilización en Aprendizaje Automático, ya que permite la configuración de algoritmos, mediante el ajuste de hiperparámetros, sin la intervención humana.”



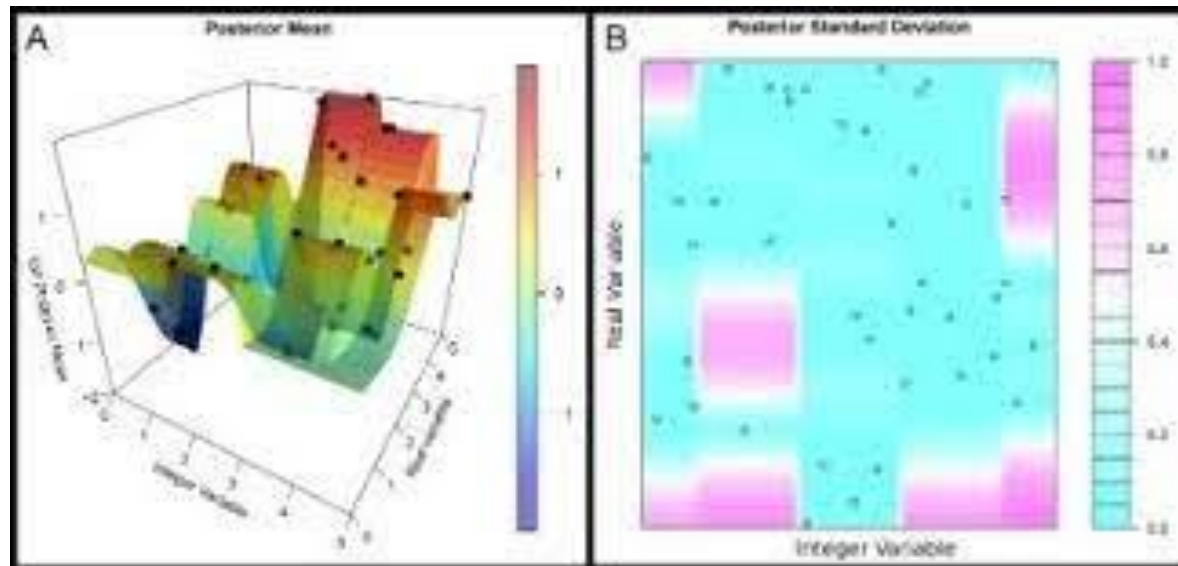
## Optimización bayesiana

“Se trata de una herramienta utilizada para resolver problemas de optimización global sobre funciones objetivo de caja negra, esto es, que no tienen expresión analítica conocida, con evaluaciones muy costosas y cuyas observaciones pueden ser ruidosas. Para ello, hace uso de un modelo probabilístico (típicamente un Proceso Gaussiano) que representa la incertidumbre sobre el objetivo, y construye una función de adquisición de forma iterativa sobre el mismo, que son criterios de utilidad que deciden los puntos a evaluar en la búsqueda del óptimo.”

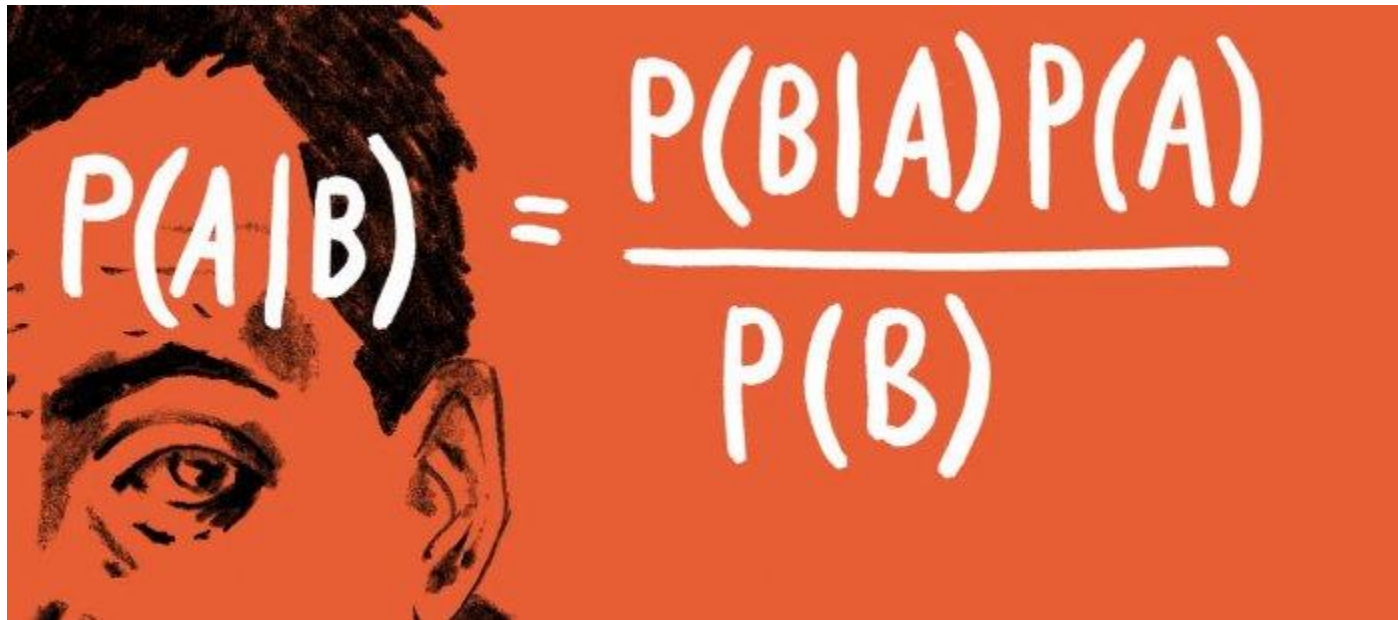


## Optimización bayesiana

La optimización bayesiana es un enfoque para encontrar la configuración óptima de parámetros de un sistema, generalmente cuando la evaluación de cada configuración es costosa o lleva mucho tiempo. Este enfoque utiliza el teorema de Bayes para construir un modelo probabilístico de la función objetivo que se está optimizando y luego utiliza este modelo para tomar decisiones informadas sobre qué configuración probar a continuación.



## Optimización bayesiana


$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

# Optimización bayesiana

- El uso de variables aleatorias, o, más en general, de cantidades desconocidas, para modelar todas las fuentes de incertidumbre en los modelos estadísticos.
- La necesidad de determinar la distribución de probabilidad previa, teniendo en cuenta la información disponible (antes).
- El uso secuencial del teorema de Bayes: cuando se disponga de más datos, calcular la distribución posterior utilizando la fórmula Bayes; posteriormente, la distribución posterior se convierte en el siguiente antes.
- Para el frecuentista una hipótesis es una proposición (que debe ser verdadera o falsa), por lo que la probabilidad frecuentista de una hipótesis es uno o cero. En la estadística bayesiana, una probabilidad asignada a una hipótesis puede diferir de 0 o 1 si el valor de verdad es incierto.



# Optimización bayesiana

- **Definición del Espacio de Parámetros:** Se define un espacio de búsqueda que describe todas las posibles configuraciones de parámetros que se pueden evaluar.
- **Evaluación de la Función Objetivo:** Se evalúa la función objetivo para una selección inicial de configuraciones de parámetros.
- **Construcción del Modelo Probabilístico:** Se utiliza el teorema de Bayes para construir un modelo probabilístico de la función objetivo. Este modelo captura la relación entre las configuraciones de parámetros y las observaciones de la función objetivo.
- **Selección de la Próxima Configuración Para Evaluar:** Con el modelo probabilístico en su lugar, se utiliza un criterio de selección (como la incertidumbre de la predicción o la esperanza de mejora) para decidir cuál configuración de parámetros probar a continuación.
- **Evaluación de la Nueva Configuración:** Se evalúa la función objetivo para la nueva configuración de parámetros seleccionada.
- **Actualización del Modelo Probabilístico:** Con la nueva observación, se actualiza el modelo probabilístico para reflejar la información adicional.
- **Iteración:** Se repiten los pasos 4-6 hasta que se alcance un criterio de detención, como un número máximo de evaluaciones.

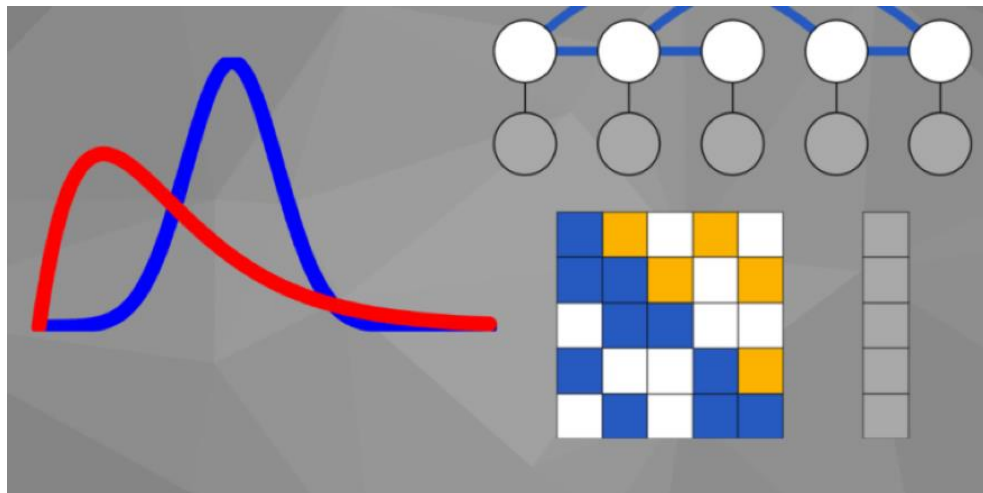
# Optimización bayesiana (valores esperados)

$$E\{g(X, Y)|Y = y_j\} = \sum_{i=1}^n g(x_i, y_j)P(X = x_i|Y = y_j),$$
$$E\{g(X, Y)|X = x_i\} = \sum_{j=1}^m g(x_i, y_j)P(Y = y_j|X = x_i).$$

<https://optuna.readthedocs.io/en/stable/tutorial/index.html>

## Optimización bayesiana (valores esperados)

La optimización bayesiana es particularmente útil en problemas donde el espacio de parámetros es grande y la evaluación de la función objetivo es costosa, como en la optimización de hiperparámetros de modelos de aprendizaje automático o en el diseño de experimentos en ciencia e ingeniería. Al aprovechar la información probabilística y tomar decisiones inteligentes sobre qué configuraciones probar a continuación, la optimización bayesiana puede converger rápidamente hacia la configuración óptima del sistema.



# Optimización bayesiana (proceso estocastico)

Un proceso estocástico es un concepto matemático que sirve para representar magnitudes aleatorias que varían con el tiempo o para caracterizar una sucesión de variables aleatorias (estocásticas) que evolucionan en función de otra variable, generalmente el tiempo. Cada una de las variables aleatorias del proceso tiene su propia función de distribución de probabilidad y pueden o no estar correlacionadas entre sí.



## Optuna

Es una biblioteca de optimización de hiperparámetros automatizada que permite encontrar los mejores hiperparámetros para un modelo de aprendizaje automático de manera eficiente. Funciona mediante la búsqueda sistemática de los hiperparámetros que minimizan o maximizan una función objetivo definida por el usuario.



OPTUNA

## Optuna

- **Definir la función objetivo:** Definir una función objetivo que evalúa el rendimiento del modelo con respecto a los hiperparámetros. Esta función debe tomar los hiperparámetros como argumentos y devolver una puntuación de métrica.
- **Crear un estudio:** Un estudio es el objeto principal de Optuna que gestiona la optimización de hiperparámetros. Se puede crear un estudio utilizando la función `optuna.create_study()`.
- **Ejecutar la optimización:** Una vez creado el estudio, se puede ejecutar la optimización llamando al método `optimize()` y pasando la función objetivo y el número de iteraciones (`n_trials`) que deseas realizar.

Optuna explora el espacio de búsqueda de hiperparámetros utilizando diferentes algoritmos de optimización, como TPE (Tree-structured Parzen Estimator) o Random Search, para encontrar los mejores hiperparámetros que minimizan o maximizan la función objetivo.