

# Trabajo 4 – Expectativas y Dinámica Macroeconómica

## Introducción a ML con Python

Abad Anguila. Robledo Meyli. Pizarro Alberto. Navarro Damaris.

Universidad de Piura

28 de Noviembre, 2025

# Contenido

- 1 Introducción
- 2 Trabajo 1: EDA
- 3 Trabajo 2: Modelo Logit
- 4 Trabajo 3: Modelos No Lineales
- 5 Trabajo 4: Causalidad y Redes Neuronales
- 6 Conclusiones

# Introducción: Pregunta de Investigación

**¿En qué medida la dinámica macroeconómica —inflación, tipo de cambio, actividad económica y política monetaria— influye en las expectativas económicas?**

- Relevancia en consumo, inversión y decisiones empresariales.
- Perú ha tenido choques recientes: pandemia, inflación, volatilidad.
- Se busca relacionar indicadores macro con optimismo/pesimismo.

# Introducción: Enfoque Metodológico

- ① EDA para caracterizar datos y detectar patrones macro.
- ② Modelo Logit como línea base.
- ③ Modelos no lineales: Random Forest y XGBoost.
- ④ Análisis causal mediante DAG.
- ⑤ Modelo de red neuronal MLP.

# Trabajo 1: Distribución de Expectativas

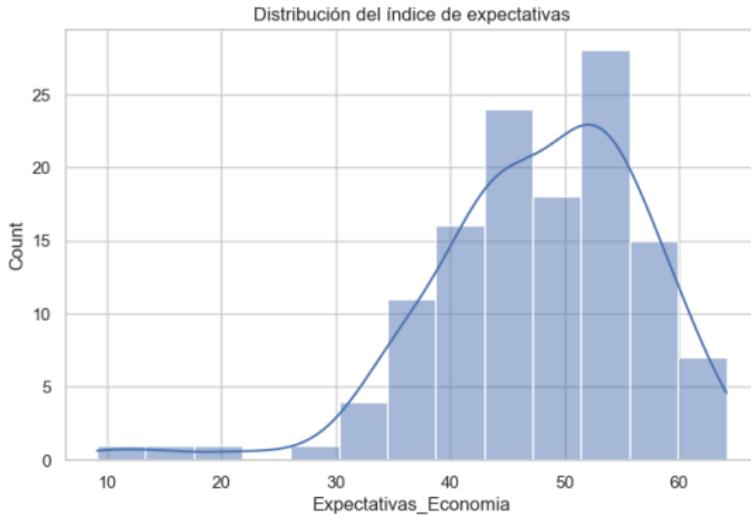


Figure: Distribución del índice de expectativas.

- Valores concentrados entre 40–55.
- Picos de pesimismo moderado (choques económicos).
- Expectativas\_dummy: 45% optimistas vs 55% no optimistas.

# Trabajo 1: Variable Objetivo Binaria

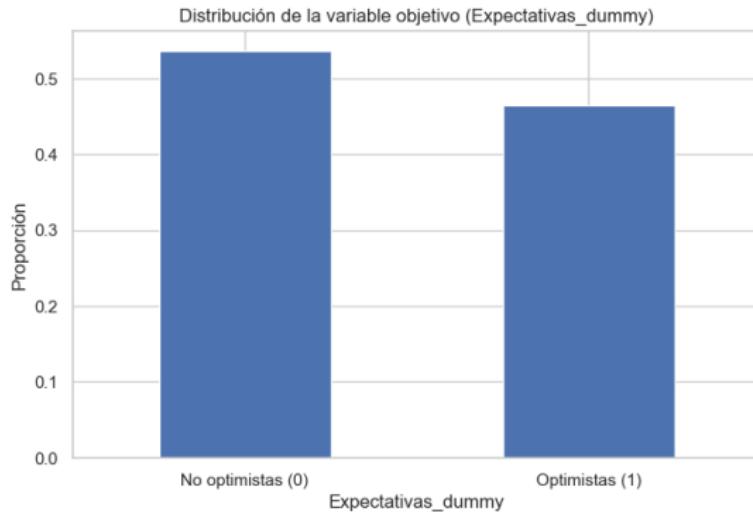


Figure: Proporción de optimistas vs no optimistas.

- La proporción de clases es relativamente balanceada: 55% “no optimistas” y 45% “optimistas”.
- No se requieren técnicas de oversampling/undersampling.
- El umbral de 50 para expectativas es razonable para representar cambios económicos.

# Trabajo 1: Series Macroeconómicas

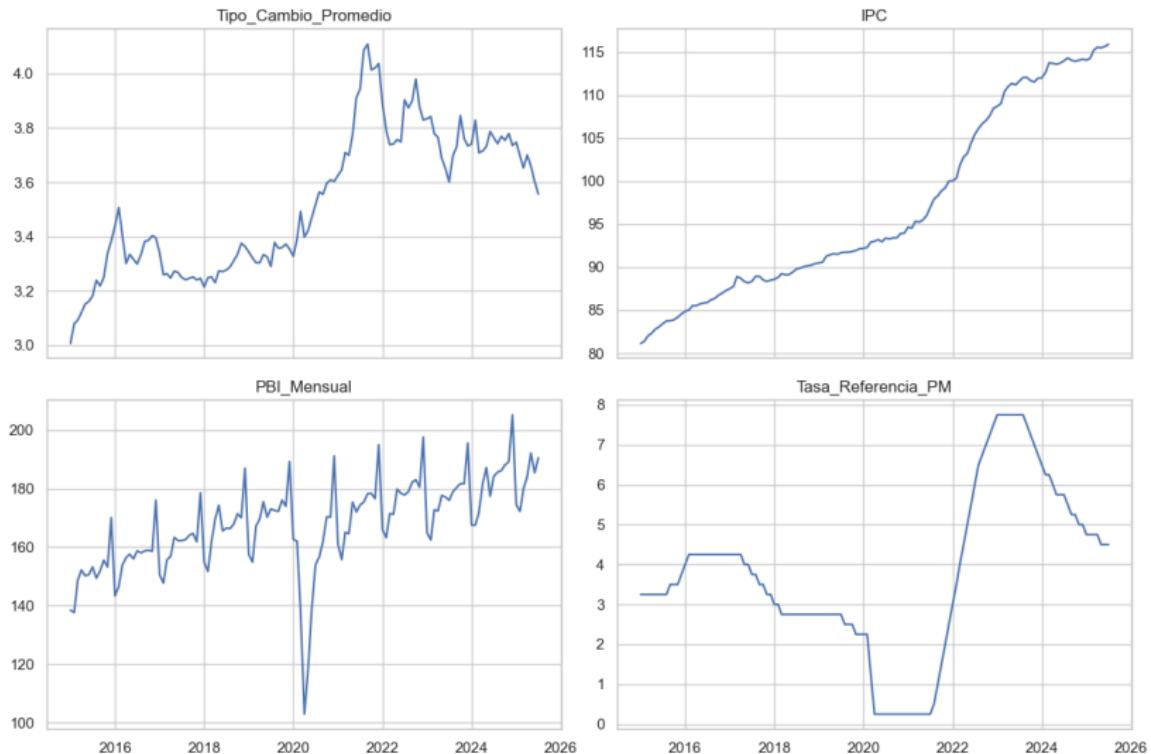


Figure: Tipo de cambio, IPC, PBI mensual y tasa de referencia.

# Trabajo 1: Correlación Macro–Expectativas

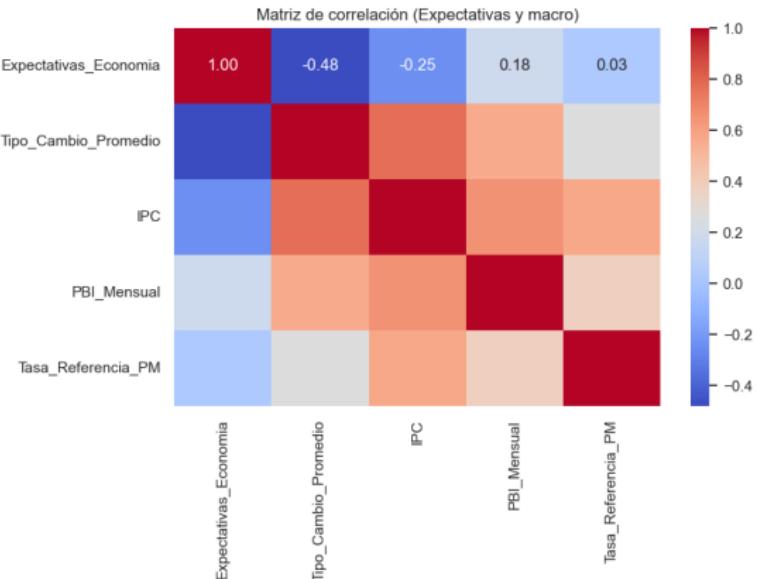


Figure: Matriz de correlación.

- Tipo de cambio vs expectativas: correlación negativa moderada.
- IPC: relación negativa leve.
- PBI: relación positiva débil.

# Trabajo 2: Feature Engineering

## 1. Rezagos (Lag 1)

- Tipo de cambio, IPC, PBI y tasa de referencia.
- Capturan efectos dinámicos y evitan simultaneidad.

## 2. Variaciones mensuales

$$\Delta X_t = X_t - X_{t-1}$$

- Permiten identificar shocks y cambios recientes.

## 3. Interacciones

TC × IPC: Tipo de cambio vs Inflación.

TC × Tasa: Política monetaria y tipo de cambio.

## 4. Transformaciones logarítmicas de TC, PBI e IPC

- Reducen heteroscedasticidad.
- Facilitan interpretación como elasticidades.

## Trabajo 2: PCA - Resultados

El PCA resume toda la información de 17 variables macroeconómicas en solo dos componentes:

- PC1 explica 54% de la variabilidad total. Captura la posición macroeconómica general: inflación, tipo de cambio y tasa suelen moverse juntas.
- (PC2) explica 14%. Captura fluctuaciones más específicas: ciclos del PBI o movimientos aislados del tipo de cambio.
- Varianza acumulada: 67.46%. Con solo dos ejes ya se puede visualizar dos tercios de toda la complejidad del conjunto macroeconómico, lo que es bastante razonable para una economía real.
- Interpretación económica: El PCA muestra que la economía peruana tiene un núcleo macro “común”, pero también shocks secundarios más específicos. Esto es coherente con los ciclos económicos recientes: pandemia, inflación global, depreciación del sol, etc.

# Trabajo 2: Biplot PCA simple coloreado por la clase

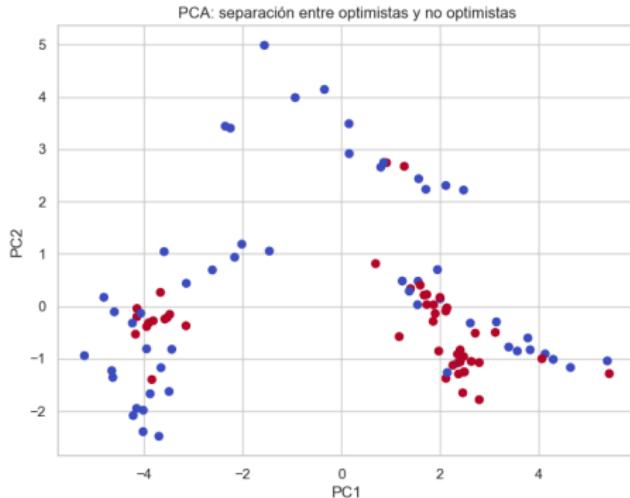


Figure: PCA: dos componentes principales.

Se observan zonas donde un grupo domina:

Cuando PC1 toma valores altos (zona derecha), se concentran más no optimistas en períodos de inflación y tipo de cambio altos.

En zonas donde PC1 es bajo (izquierda), aparecen más optimistas, característicos de períodos macro más estables.

## Trabajo 2: Modelo de Regresión Logística

Transformamos nuestra variable de Expectativas en una variable categórica, donde:

- 1 = optimismo si  $> 50$
- 0 = Pesimismo si  $< 50$

Cuando la variable objetivo es categórica, se aplicó un modelo logístico:

$$P(ExpEconómicas = 1|X) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 TC + \beta_2 TPM + \beta_3 PBI + \beta_4 IPC)}}$$

## Trabajo 2: Resultados del Logit

- Accuracy: Logra predecir correctamente cerca del 77% de los casos.
- Precision/Recall/F1 = 0.75. El modelo es capaz de capturar al 75% de los agentes realmente optimistas.
- ROC-AUC = 0.79, hay un 79% de probabilidad de que el modelo distinga correctamente entre un optimista y un no optimista usando las variables macro.
- Capta bastante bien el vínculo entre la macro y las expectativas, pero deja espacio para mejoras con modelos no lineales. Es un benchmark sólido que justifica comparar con modelos más complejos.

# Trabajo 3: Random Forest

- La profundidad óptima indica que las reglas de decisión no necesitan ser demasiado profundas. Sugiere que algunos indicadores macro clave ya contienen suficiente información para separar optimistas de no optimistas.
- El uso de 200 árboles estabiliza muy bien las predicciones y evita que el modelo dependa de un solo patrón temporal.
- AUC 0.92. El modelo distingue muy bien los dos tipos de agentes bajo distintos escenarios macroeconómicos.
- Cuando la macro se deteriora, el RF detecta estos patrones y los transforma en alta probabilidad de pesimismo. Cuando hay estabilidad monetaria y mayor actividad, el modelo lo identifica como señales de optimismo.

# Trabajo 3: XGBoost

- El learning\_rate=0.01 muestra que las relaciones macro-expectativas son delicadas; el modelo avanza “paso a paso” para no sobreajustarse.
- Una profundidad de 3 niveles indica que los patrones relevantes no necesitan árboles muy complejos, sino reglas macroeconómicas simples, pero bien combinadas.
- La similitud en desempeño con Random Forest (AUC 0.91) refuerza la consistencia del resultado.
- Los agentes parecen reaccionar no solo a niveles de inflación o tipo de cambio, sino también a su combinación con rezagos, variaciones y señales de política monetaria. XGBoost es especialmente bueno detectando “shocks” recientes.

# Trabajo 3: Comparación de Modelos

```
==== ROC-AUC POR MODELO ====
Logit : 0.7916666666666667
Random Forest : 0.9166666666666666
XGBoost : 0.9136904761904763
```

Figure: Comparación de modelos

- El Logit capta la relación lineal básica entre la macro y las expectativas → funciona como referencia inicial.
- Random Forest y XGBoost superan ampliamente al Logit → indican que la relación real NO ES lineal, sino moderadamente compleja.
- Las métricas de AUC cercanas a 0.92 muestran que los modelos entienden bien cómo cambios en inflación, tipo de cambio, PBI y tasa de referencia influyen en el estado de ánimo de consumidores y empresas.

# Trabajo 3: Costo según Umbral

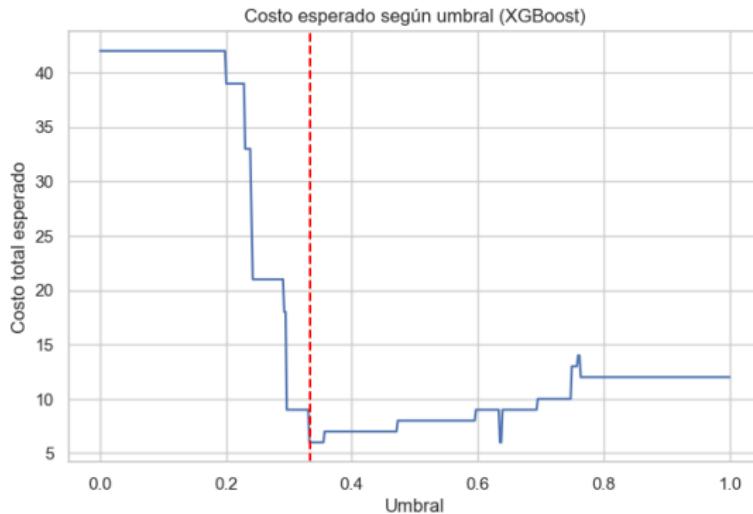


Figure: Costo esperado al variar el umbral.

- El costo de falsos positivos se definió como más alto, sobreestimar el optimismo es más peligroso que subestimar.
- El modelo aprende que, en la economía peruana, los shocks negativos son más frecuentes o más severos, por lo que es racional que el umbral óptimo se desplace hacia valores más conservadores.

# Trabajo 4: DAG Causal

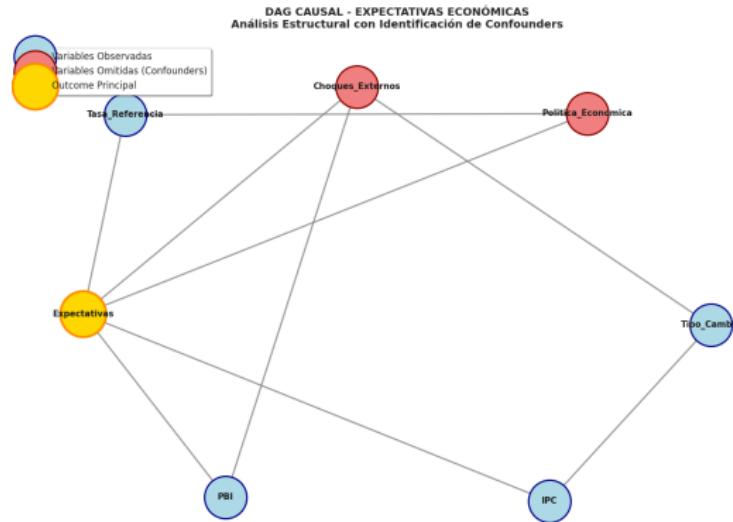


Figure: DAG causal de expectativas económicas.

Las expectativas económicas se forman mediante un mecanismo causal multivariado donde el crecimiento (PBI), la inflación (IPC) y la tasa de interés afectan directamente la percepción de los agentes sobre el ciclo económico.

Los choques externos y la política económica operan como confounders clave, distorsionando las relaciones observables al impactar simultáneamente múltiples variables.

# Trabajo 4: MLP CLASSIFIER (RED NEURONA)

- El MLP obtuvo un buen desempeño con 88.5% de accuracy y un ROC-AUC destacado de 0.952, superando a los modelos anteriores.
- Esto confirma que la arquitectura de red neuronal con capas (32,16) y optimizador Adam fue efectiva para capturar las relaciones no lineales complejas en los datos macroeconómicos.
- El alto recall (91.7%) podría indicar que el modelo es particularmente bueno identificando casos optimistas, mientras que la precisión (84.6%) sugiere cierta tendencia a falsos positivos.

# Trabajo 4: COMPARACIÓN FINAL: LOGIT vs RF vs XGB vs MLP

- Podemos decir que el MLP fue el que mejor resultado que obtuvo (95.2% de ROC-AUC), seguido muy de cerca por Random Forest (91.7%) y XGBoost (91.4%). El modelo Logit tuvo el desempeño más bajo (79.2)
- Los modelos más complejos (MLP, Random Forest, XGBoost) funcionan mejor que el modelo simple (Logit)
- El MLP, siendo una red neuronal, captura mejor las relaciones complejas en los datos económicos, sin necesidad de un ajuste muy detallado.

# Conclusiones

- ① El EDA muestra que las expectativas económicas varían de forma consistente con los ciclos macro.
- ② El modelo Logit sirve como línea base: capta parte de la señal, pero no aprovecha todas las no linealidades.
- ③ Random Forest y XGBoost mejoran claramente el desempeño (ROC-AUC 0.9), lo que sugiere relaciones no lineales y efectos de interacción entre variables macro.
- ④ El MLP obtiene un desempeño similar/comparable, mostrando que una arquitectura de red neuronal moderadamente compleja también puede capturar patrones útiles.
- ⑤ El DAG explica los supuestos causales: los choques externos y la política económica actúan como confounders al afectar tanto las variables macro como las expectativas.

¡Gracias!