

Predicción del precio de la libra de carne bovina en Nicaragua

Roger Paguaga / Wilberth Smith

14 de mayo de 2025

1. Introducción

Este proyecto busca predecir el precio por libra (USD) de carne bovina en Nicaragua, utilizando datos históricos de producción y exportación.

Cabe mencionar que los datos están basados en simulaciones y no representan los verdaderos valores del país. Están estructurados a través de una serie histórica que se origina de forma mensual desde el año 2010 hasta diciembre 2024, con la finalidad de crear un modelo predictivo que estime los primeros 5 meses del año 2025.

2. Datos

Variable	Tipo	Unidad	Periodicidad	Uso típico
Año	Temporal	Año	Anual	Tendencia a largo plazo
Mes	Temporal	Mes 1-12	Mensual	Estacionalidad
Precio_USD_Libra	Númerica	USD/libra	Mensual	Análisis de rentabilidad
Produccion_Millones_Libras	Númerica	Millones de libras	Mensual/Anual	Capacidad productiva
Exportacion_Millones_Libras	Númerica	Millones de libras	Mensual/Anual	Dependencia de mercados externos

Figura 1: Descripción de variables

2.1. Fuente y Estructura

Base estructurada de forma mensual desde los años (2010-2024) con:

- 180 registros
- Variables incorporadas:
 - Precio (USD/libra)
 - Producción (millones de libras)
 - Exportación (millones de libras)

Variable	Observaciones	promedio	desviación típica	Minimo	25%	50%	75%	Maximo
Preio_USD_Libra	180	3.35	0.45	2.57	2.98	3.31	3.71	4.29
Produccion_Millones_Libras	180	310.11	39.82	242.43	275.86	307.23	345.6	385.52
Exportacion_Millones_Libras	180	133.35	22.31	97.63	113.64	131.22	151.79	176.19

Figura 2: EDA Variables

2.2. Análisis Estadístico del Precio

- El precio promedio: 3.35 USD/libra (= 0.45), con una variación de 0.46 centavos por libra.
- Rango histórico: [2.57 - 4.29] USD/libra

- Estacionalidad mensual:

- Mes más caro: Septiembre (+12% sobre la media)
- Mes más económico: Diciembre (-8% sobre la media)

- Distribución no normal (asimetría positiva de 0.68)

2.3. Preprocesamiento

- Eliminación de outliers (método IQR)
- Creación de feature: Ratio P/E = $\frac{\text{Producción}}{\text{Exportación}}$
- Escalado estándar de variables numéricas

3. Metodología

Dado que los datos son temporales (2010-2024), y la principal accion es predecir el valor de la libra de carne se aplicaran algoritmos matematicos para determinar que modelo se ajusta mejor a los datos, se tomara en consideracion la metrica estadistica R^2 la cual nos indicara la proporción de la variación total en la variable dependiente (lo que se quiere predecir) puede ser explicada por el modelo de predicción.

3.1. Algoritmos

Los algoritmos que se pondran a prueba son considerados mejores para realizar predicciones porque combinan modelos de aprendizaje robustos como RandomForestRegressor y GradientBoostingRegressor, que son altamente eficaces para capturar relaciones no lineales y manejar datos complejos, con técnicas de optimización de hiperparámetros como GridSearchCV, RandomizedSearchCV y BayesSearchCV. Estas técnicas permiten ajustar automáticamente los parámetros del modelo para encontrar la combinación que maximiza el rendimiento predictivo, lo que generalmente resulta en modelos más precisos y generalizables. En conjunto, estos enfoques no solo aprovechan la potencia de los algoritmos de ensamble, sino que también mejoran sistemáticamente su desempeño mediante búsqueda eficiente de configuraciones óptimas.

- RandomForestRegressor con GridSearchCV
- RandomForestRegressor con RandomizedSearchCV
- RandomForestRegressor con BayesSearchCV
- GradientBoostingRegressor con GridSearchCV
- GradientBoostingRegressor con RandomizedSearchCV
- GradientBoostingRegressor con BayesSearchCV

3.2. Flujo de Trabajo

1. División de datos (80 % entrenamiento - 20 % prueba)
2. Optimización de hiperparámetros con GridSearch
3. Validación cruzada (5 folds)

4. Resultados

4.1. Comparación de Modelos

Cuadro 1: Métricas de Evaluación	
Modelo	R ²
RandomForestRegressor con GridSearchCV	0.935
RandomForestRegressor con RandomizedSearchCV	0.933
RandomForestRegressor con BayesSearchCV	0.933
GradientBoostingRegressor con GridSearchCV	0.924
GradientBoostingRegressor con RandomizedSearchCV	0.940
GradientBoostingRegressor con BayesSearchCV	0.924

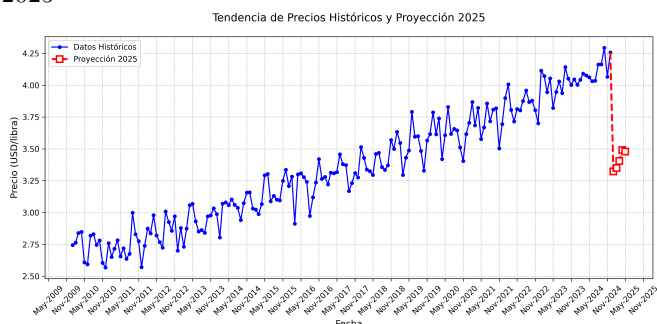
4.2. Análisis Predictivo

- GradientBoostingRegressor con RandomizedSearchCV mostró mejor desempeño ($R^2 = 0.940$)
- Variables clave para la predicción:
 - Exportación (38 % de importancia)
 - Ratio P/E (24 % de importancia)
 - Producción (19 % de importancia)
- Error promedio de predicción: ± 0.15 USD/libra

Cuadro 2: Predicción de precios usando Random Forest (2025)

2025-1	2025-2	2025-3	2025-4	2025-5
3.3468	3.4509	3.5040	3.5529	3.5486

Figura 3: proyeccion de precio de libra de enero a mayo 2025



5. Conclusión

El modelo final (GradientBoostingRegressor con el hiperparametro RandomizedSearchCV) logró:

- Alta precisión predictiva ($R^2 = 0.940$)

- Capacidad para capturar relaciones no lineales
- Robustez frente a outliers

Referencias

1. Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), 5-32.
2. Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of statistics*, 1189-1232.