Pronosticando inflaciones de canastas de alimentos desagregadas en Colombia usando un modelo XGBoost

Borradores de Economia
Número:
1335
DOI:
https://doi.org/10.32468/be.1335
Publicado:
Miércoles, 10 Diciembre 2025
Authors:
Cesar Anzola-Bravo[a],
Poveda-Olarte Paola[a]
Ver más
[a]Banco de la República, Colombia
Clasificación JEL:
C53, E31, E37
Palabras clave:
pronóstico macroeconómico, Inflacion de alimentos, Machine Learning
Resumen:

Los precios de los alimentos han sido uno de los principales factores que contribuyen a la inflación en Colombia. Estos son particularmente sensibles a factores externos como choques climáticos, interrupciones en las cadenas globales de valor y choques en los precios de los productos básicos a nivel global, lo que resulta en fluctuaciones impredecibles de precios. Este documento tiene dos objetivos. En primer lugar, busca estimar y evaluar métodos para pronosticar 33 canastas homogéneas de inflación de alimentos, ofreciendo herramientas que puedan ayudar a los hacedores de política anticipar los factores que afectan la inflación de alimentos futura. Esto incluye tanto modelos tradicionales de series de tiempo como enfoques modernos de machine learning. En segundo lugar, se propone mejorar la interpretabilidad de las predicciones de los modelos mediante técnicas de explainableAI. Para ello, proponemos un algoritmo de selección de variables que identifique las variables explicativas más relevantes, y utilizamos valores SHAP (SHapley Additive exPlanations) para cuantificar la contribución de cada variable explicativa en las predicciones del modelo. Nuestros hallazgos indican que los modelos de machine learning superan a los enfoques tradicionales en el pronóstico de la inflación de alimentos, logrando una mayor precisión tanto en la mayoría de las canastas individuales como en la inflación de alimentos agregada.

Descargar documento

- Enfoque
- Contribución
- Resultados

Lo más reciente

[Settlement liquidity in the Colombian large-value payment system: the role of reserve requirements](#)
Fredy Gamboa-Estrada, Freddy Cepeda-López, Javier Iván Miguélez-Márquez
[Assessing the effects of recent provisioning rules on consumer credit allocation in Colombia](#)
Fredy Gamboa-Estrada, Diego Cuesta, Camilo Sanchez-Quinto
[Geopolitical risk and emerging market sovereign risk premia](#)
Fredy Gamboa-Estrada, José Vicente Romero
[Otras Publicaciones](#)

Enfoque

Este artículo tiene como objetivo desarrollar modelos estadísticos para pronosticar la inflación mensual de los próximos 12 meses para 33 canastas que conforman el Índice de Precios al Consumidor (IPC) de alimentos. Para ello, se emplean tanto modelos tradicionales de series de tiempo como enfoques basados en Machine Learning. Cada canasta se modela de manera independiente, incorporando 4 grupos de variables explicativas relevantes para la oferta de alimentos: variables climáticas, la tasa de cambio nominal, precios de materias primas, y costos de transporte y energía. Los precios de las materias primas y la energía influyen en la inflación de alimentos al afectar los costos de producción, transporte y procesamiento. Además, el clima impacta la producción agrícola al alterar el crecimiento de cultivos, la calidad del suelo y la salud del ganado. Por otro lado, el documento busca interpretar los pronósticos mediante el uso de valores SHAP (SHapley Additive exPlanations), una herramienta ampliamente utilizada para explicar las predicciones de modelos de machine learning.

Contribución

Esta investigación contribuye a la literatura de pronósticos de inflación de alimentos al probar que los modelos basados en arboles de decisión tipo XGBoost realizan pronósticos más acertados que modelos de series de tiempo lineales. Por otro lado, mostramos que es posible descomponer las predicciones de este modelo en contribuciones de variables explicativas fundamentales. Las dinámicas de cada canasta de alimentos responden a variables como el clima, precios de materias primas, costos de transporte y el comportamiento de la tasa de cambio.

La interpretación de las predicciones de estos modelos es lograda gracias a un proceso exhaustivo de selección de las variables explicativas y a un algoritmo que proponemos para seleccionar los rezagos óptimos de las variables explicativas. Este algoritmo permite reducir el número de variables, lo que simplifica la interpretación y reduce costos computacionales.

> ***Los modelos tipo XGBoost son más precisos pronosticando la inflación de alimentos que los modelos lineales para la mayoría de las 33 canastas estudiadas, especialmente para horizontes de pronóstico más lejanos.***

Resultados

Los modelos tipo XGBoost son más precisos pronosticando la inflación de alimentos que los modelos lineales para la mayoría de las 33 canastas estudiadas, especialmente para horizontes de pronóstico más lejanos. Los errores de pronóstico del modelo XGBoost fueron entre un 5% y un 60% menores que los de los modelos lineales, dependiendo de la canasta y del horizonte de pronóstico, y para la canasta agregada de inflación los errores fueron en promedio 25% menores.

Existe un alto grado de heterogeneidad en cuanto a la explicación de los pronósticos dependiendo de la canasta a estudiar. Para ciertos grupos de alimentos, como los alimentos perecederos, el clima y la persistencia de la inflación son factores relevantes. En contraste, otros alimentos, como los industriales, son principalmente explicados por costos de materias primas y persistencia de la inflación.