



El poder predictivo de Gemini: Un nuevo horizonte para predecir la inflación.

Germán Camilo Vásquez Herrera

Director: Juan Camilo Sosa Martínez

Tutor: Aarón Levi Garavito Acosta

Departamento de Estadística - Universidad Nacional de Colombia Sede Bogotá



Resumen

En este estudio, se compararon tres métodos para pronosticar la inflación en Colombia: el modelo de IA Gemini de Google y las encuestas del Banco de la República (mensual a analistas del mercado y trimestral a empresarios colombianos). Gemini demostró ser el más preciso para el largo plazo (4-12 pasos adelante), pese a desafíos técnicos como claves API y formatos de respuesta. La encuesta mensual del Banco fue superior en el corto plazo (0-3 pasos adelante), mientras que la encuesta trimestral a empresarios tuvo un desempeño deficiente. Estos resultados resaltan el potencial de la IA en proyecciones económicas a largo plazo, subrayando la necesidad de gestionar sus retos técnicos.

Objetivos

- Automatizar consultas diarias a la IA Gemini.
- Comparar la precisión predictiva de Gemini con la encuesta mensual a analistas de mercado y la encuesta trimestral a empresarios colombianos.
- Identificar limitaciones técnicas y computacionales en la predicción.
- Sugerir mejoras futuras a partir de los hallazgos.

Metodología

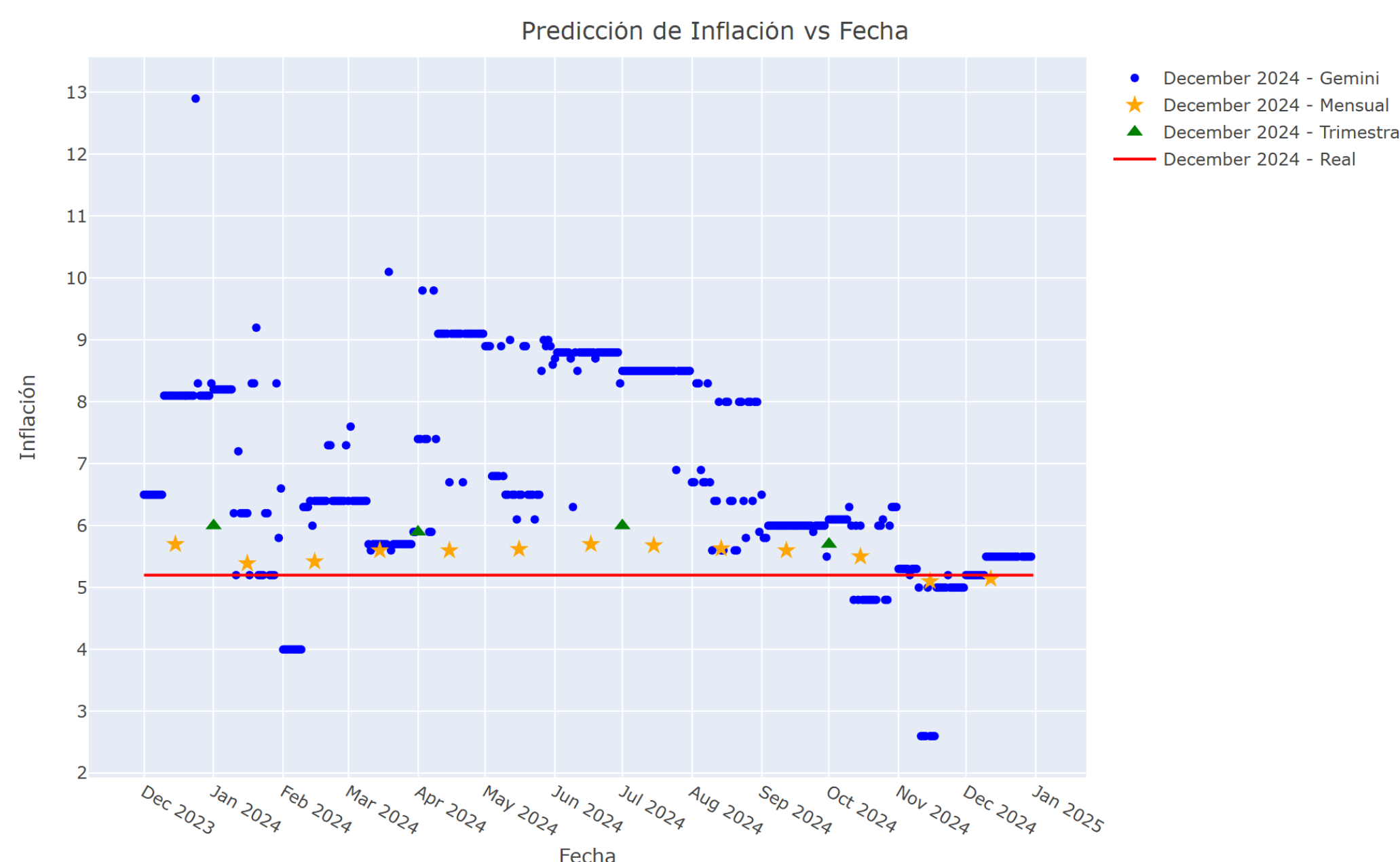
Este informe compara tres métodos de pronóstico de inflación en Colombia: la **IA Gemini de Google** y dos encuestas del **Banco de la República** (mensual a analistas del mercado y trimestral a empresarios colombianos). Gemini predice a 12 meses usando Python y R, mientras que las encuestas cubren plazos desde el mes actual hasta 11 o 12 meses adelante. La precisión de cada método se evalúa mediante **RMSE** y **MAE**.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}, \quad MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

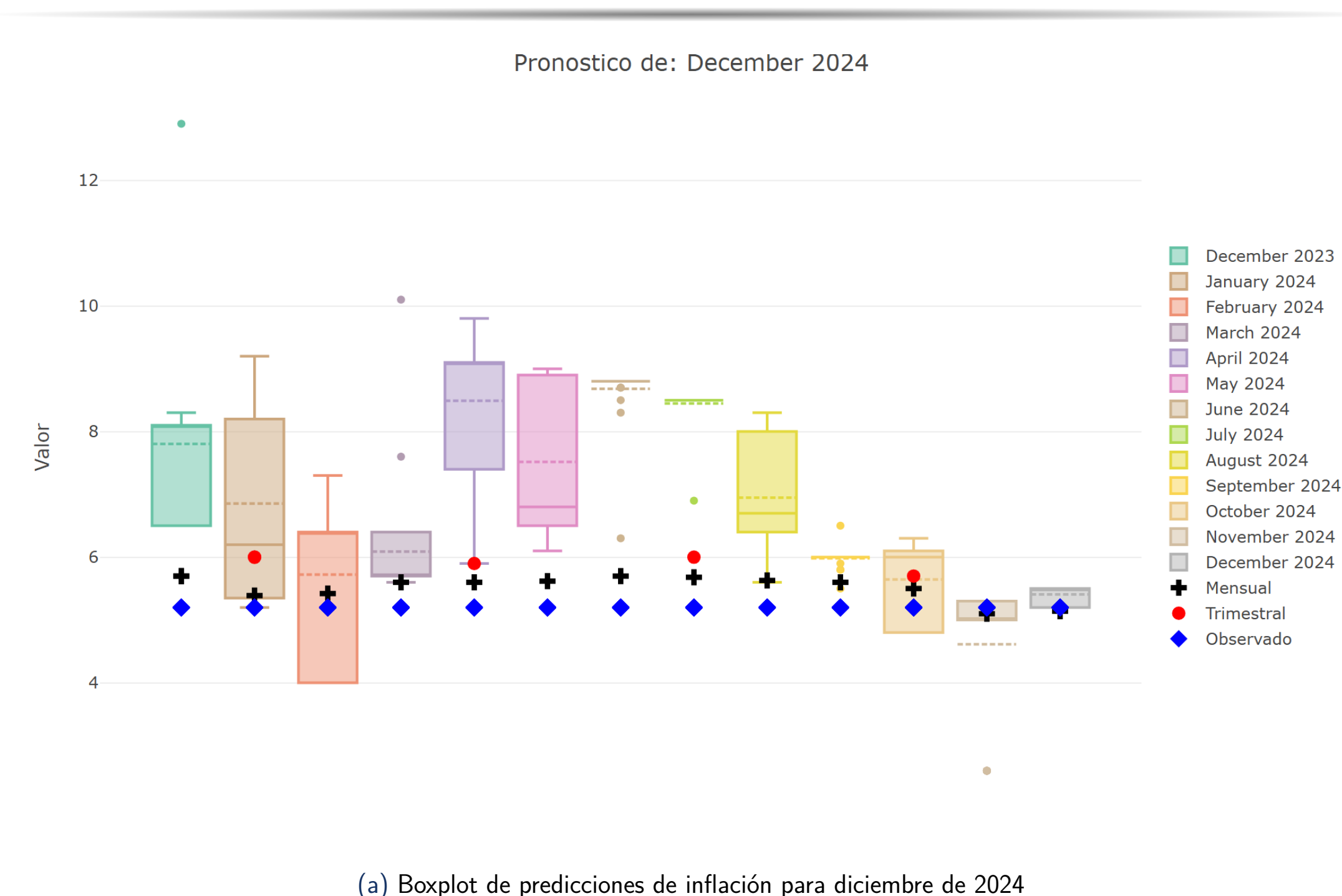
donde y_i es el valor real, \hat{y}_i es el valor predicho y n es el número de observaciones. Estas métricas posibilitan evaluar la **exactitud** y **consistencia** de los métodos, determinando cuál genera los **pronósticos inflacionarios más fiables**.

Análisis y Resultados

En esta sección, se evalúa la precisión predictiva de cada método de pronóstico de inflación usando **RMSE** y **MAE**. Se analizan las **predicciones a 12 pasos adelante** generadas mensualmente entre **enero de 2022** y **marzo de 2025**, considerando las limitaciones del Banco en ofrecer proyecciones completas. Posteriormente, se comparan los pronósticos de los tres métodos con los valores de inflación observados.



(a) Dispersión de predicciones de inflación para diciembre de 2024



(a) Boxplot de predicciones de inflación para diciembre de 2024

Inicialmente, la **encuesta trimestral** tuvo un desempeño deficiente, pero mejoró al final del estudio, sin igualar la precisión de la **IA Gemini** o la **encuesta mensual**, las más confiables.

Error Cuadrático Medio (RMSE)							
Paso	G.M.M.	G.T.M.	G.M.Me.	G.T.Me.	G.Tót.	B.M.	B.T.
0	0.819	0.797	0.906	0.776	0.943	1.267	
1	0.957	0.935	1.130	1.002	1.188	0.541	
2	1.193	1.156	1.353	1.268	1.420	0.758	1.778
3	1.421	1.378	1.565	1.503	1.660	1.065	
4	1.631	1.587	1.766	1.714	1.874	1.883	
5	1.835	1.794	1.955	1.918	2.092	2.277	2.553
6	2.005	1.970	2.100	2.105	2.296	2.619	
7	2.136	2.114	2.216	2.232	2.465	2.852	
8	2.220	2.198	2.306	2.356	2.600	3.541	3.445
9	2.298	2.285	2.394	2.474	2.736	3.879	
10	2.353	2.346	2.462	2.551	2.859	4.443	
11	2.403	2.405	2.553	2.649	2.953	5.005	4.028
12	2.316	2.345	2.565	2.605	2.782	4.278	

Tabla: Resultados RMSE

Durante enero de 2022 y marzo de 2025, **Gemini** (promediando 20 días) fue el más preciso a **largo plazo** (4-12 pasos), con los errores más bajos (RMSE y MAE). La **encuesta mensual** del Banco destacó en el **corto plazo** (0-3 pasos), mientras que la **encuesta trimestral** fue la menos efectiva. Esto posiciona a Gemini como el mejor predictor a largo plazo.

Conclusiones

- Se implementó un **sistema semiautomático** para Gemini, que requiere **actualizaciones periódicas** por limitaciones técnicas.
- Gemini** fue más **preciso** (menores RMSE y MAE) a **largo plazo** (4-12 pasos), mientras la encuesta mensual **destacó** a **corto plazo** (0-3 pasos), y la trimestral tuvo un **pésimo desempeño**.
- Las limitaciones incluyeron **caducidad** y **cupó** de claves API, además de **formatos variables** en las respuestas de Gemini, afectando la automatización.
- Se sugieren mejoras como un manejo **robusto** de claves API, procesamiento de datos **flexible**, y explorar otras **IA** para optimizar el proceso.

Referencias

- Banco de la República, Encuesta de Expectativas Económicas (EME), *Banco de la República*, 2023, <https://suameca.banrep.gov.co/estadisticas-economicas/encuestas>.
- Hyndman, R. J. y Koehler, A. B., Another look at measures of forecast accuracy, *International Journal of Forecasting*, vol. 22, no. 4, pp. 679-688, 2006, <https://www.sciencedirect.com/science/article/abs/pii/S0169207006000239>.
- Pichai, S. y Hassabis, D., Introducing Gemini: our largest and most capable AI model, *Making AI more helpful for everyone*, 2023, <https://blog.google/technology/ai/google-gemini-ai/>.
- Chai, T. y Draxler, R. R., Root mean square error (RMSE) or mean absolute error (MAE)?, *Geoscientific Model Development*, vol. 7, no. 3, pp. 1247-1250, 2014, <https://gmd.copernicus.org/articles/7/1247/2014/>.