CENTRO DE INVESTIGACIÓN Y DOCENCIA ECONÓMICAS, A.C

PREDICCIÓN DE LA INFLACIÓN MEXICANA EN TIEMPOS DE INCERTIDUMBRE: UNA EVALUACIÓN DE MODELOS DE APRENDIZAJE DE MÁQUINA Y SU IMPACTO EN LA PRECISIÓN DEL PRONÓSTICO

TESINA

QUE PARA OBTENER EL TITULO DE

LICENCIADO EN ECONOMÍA

PRESENTA

ERIK ROSAS RUIZ

DIRECTOR DE LA TESINA: DR. GONZALO ALBERTO CASTAÑEDA RAMOS

CIUDAD DE MÉXICO 2023

Resumen

En respuesta a la creciente necesidad de mejorar la precisión en el pronóstico económico y tomar decisiones informadas en tiempos de incertidumbre económica en México, esta investigación se basa en la premisa de que las herramientas de aprendizaje de máquina (ML por sus siglas en inglés) pueden revolucionar la forma en que abordamos este desafío. Específicamente, se enfoca en evaluar el impacto de estos modelos en la predicción de la inflación mexicana. Los resultados preliminares de esta investigación demuestran que los modelos de *ML* superaron de manera significativa a los enfoques tradicionales, reduciendo el error cuadrático medio (MSE por sus siglas en inglés) en un promedio del 90%. Además, se destaca la eficacia de los modelos de Auto*ML* en el proceso de pronóstico, respaldando la idea de que estas herramientas pueden ser un recurso valioso para impulsar la precisión en el pronóstico económico y la toma de decisiones en política monetaria y fiscal.

Palabras Clave: Inflación, Modelos de *Machine Learning*, Incertidumbre Económica, Economía Mexicana, Herramientas de Pronóstico, Modelos de Auto*ML*, Modelo *LSTM*, Economía y Tecnología, Evaluación de Modelos, Transformación en Pronósticos Económicos.

Índice general

1. Introducción 1
2. Revisión de literatura 4
3. Descripción de los datos 6
4. Metodología 8
4.1. Modelo de referencia: autorregresivo (AR) 8
4.2. Modelos de aprendizaje de máquina 9
4.2.1. Optimización manual de hiperparámetros 9
4.2.2. Optimización automática de hiperparámetros: Auto*ML* 13
5. Resultados 16
5.1. Predicciones para 2022 17
5.1.1. INPC General 17
5.1.2. INPC Subyacente 19
5.1.3. INPC No Subyacente 20
5.2. Predicciones para 2008 21
5.2.1. INPC General 21
5.2.2. INPC Subyacente 22
5.2.3. INPC No Subyacente 23
6. Conclusión 24
7. Bibliografía 27
8. Apéndice 29
8.1. Pronósticos del modelo *LSTM* 29
8.1.1. Predicciones 2022 29
8.1.2. Predicciones 2008 35
8.2. Pronósticos del modelo Auto*ML* 37
8.2.1. Predicciones 2022 37
8.2.2. Predicciones 2008 40
8.3. Pronósticos del modelo AR 43
8.3.1. Predicciones 2022 43
8.3.2. Predicciones 2008 46

índice de tablas

Tabla 1: Catálogo de AutoGluon para series de tiempo……14

Tabla 2: MSE para el pronóstico del INPC General (2022)……..17

Tabla 3: MSE para el pronóstico del INPC Subyacente (2022)……....19

Tabla 4: MSE para el pronóstico del INPC No Subyacente (2022)……..20

Tabla 5: MSE para el pronóstico del INPC General (2008)…….....21

Tabla 6: MSE para el pronóstico del INPC Subyacente (2008)……22

Tabla 7: MSE para el pronóstico del INPC No Subyacente (2008)…….....23

índice de figuras

Figura 1: Esquema del modelo *LSTM* 13

Figura 2: Predicción del INPC General en el periodo 2005-2022 mediante el modelo *LSTM* .18

Figura 3: Predicción del INPC General en el periodo 2008 mediante el modelo Ensamble-Auto*ML* 22

Figura 4: Calibración del modelo *LSTM* para pronosticar el INPC General.....30

Figura 5: Predicción del INPC General en el periodo 2022 mediante el modelo *LSTM*.....31

Figura 6: Calibración del modelo *LSTM* para pronosticar el INPC Subyacente.....32

Figura 7: Predicción del INPC Subyacente en el periodo 2022 mediante el modelo *LSTM*.....33

Figura 8: Calibración del modelo *LSTM* para pronosticar el INPC No Subyacente 34

Figura 9: Predicción del INPC No Subyacente en el periodo 2022 mediante el modelo *LSTM* 34

Figura 10: Predicción del INPC General en el periodo 2008 mediante el modelo *LSTM*.....35

Figura 11: Predicción del INPC Subyacente en el periodo 2008 mediante el modelo *LSTM*...36

Figura 12: Predicción del INPC No Subyacente en el periodo 2008 mediante el modelo *LSTM* 37

Figura 13: Predicción del INPC General en el periodo 2022 mediante el modelo Ensamble-Auto*ML* 38
Figura 14: Predicción del INPC Subyacente en el periodo 2022 mediante el modelo ETS-Auto*ML* 39
Figura 15: Predicción del INPC No Subyacente en el periodo 2022 mediante el modelo Temporal Fusion Transformer-Auto*ML* 40
Figura 16: Predicción del INPC General en el periodo 2008 mediante el modelo Ensamble-Auto*ML* 41
Figura 17: Predicción del INPC Subyacente en el periodo 2008 mediante el modelo ETS-Auto*ML* 42
Figura 18: Predicción del INPC No Subyacente en el periodo 2008 mediante el modelo Temporal Fusion Transformer-Auto*ML* 43
Figura 19: Predicción del INPC General en el periodo 2022 mediante el modelo AR 44
Figura 20: Predicción del INPC Subyacente en el periodo 2022 mediante el modelo AR 45
Figura 21: Predicción del INPC No Subyacente en el periodo 2022 mediante el modelo AR . 46
Figura 22: Predicción del INPC General en el periodo 2008 mediante el modelo AR 47
Figura 23: Predicción del INPC Subyacente en el periodo 2008 mediante el modelo AR 48
Figura 24: Predicción del INPC No Subyacente en el periodo 2008 mediante el modelo AR . 49

Capítulo 1

Introducción

Los fenómenos económicos siguen patrones dinámicos difíciles de pronosticar, modelar o identificar de manera certera. Hasta ahora, hemos confiado en la capacidad de la mente humana y la ciencia para elaborar procesos generadores de datos para caracterizar a estos fenómenos. Ahora bien, los humanos poseemos el sistema de procesamiento más avanzado que la evolución haya concebido (cerebro); no obstante, es bastante limitado en cuanto a la capacidad de cómputo de patrones complejos.

Con la revolución de la información y el desarrollo tecnológico, hoy existen sistemas que nos permiten aproximarnos a estos fenómenos de manera más exacta, o cuando menos, con un menor margen de error. Estos sistemas a los que nos referimos pertenecen al campo de la Inteligencia Artificial y son los llamados modelos de aprendizaje de máquina (*ML* por sus siglas en inglés). A partir de este esquema, podemos tomar un caso particular de fenómeno económico complejo: la inflación. De esta manera el reto de esta tesis es elaborar un pronóstico preciso de la inflación mexicana, lo que se dificulta en periodos de incertidumbre económica. Tarea que resulta fundamental para tomar decisiones informadas y diseñar estrategias efectivas en el ámbito de la política monetaria y fiscal.

La hipótesis que esta tesina busca probar es que los modelos de aprendizaje de máquina son capaces de captar de manera más efectiva la dinámica no-lineal de la inflación durante periodos de incertidumbre económica, en comparación con los modelos convencionales basados en Autorregresión (*AR*) y evaluados en términos de error cuadrático medio (*MSE*). Para ello, evaluaremos los pronósticos en la crisis financiera de 2007-2008 y el periodo postpandemia de 2022.

Estos dos años se destacan en la historia económica reciente de México debido a eventos que generaron una considerable incertidumbre, como la crisis financiera global de 2007-2008, la pandemia de COVID-19 en 2020 y el conflicto entre Rusia y Ucrania en 2022. Es relevante destacar que los efectos de la pandemia de COVID-19 en la inflación alcanzaron su punto más

alto a finales de 2022 y principios de 2023, presentando una inflación anual cercana al 8% en la primera quincena de enero de 2023. De tal forma, el año 2022 fue elegido un año clave para probar los modelos.

En la literatura existen diversas contribuciones sobre el pronóstico de la inflación mediante modelos de *ML*. Por ejemplo, Medeiros et al. (2019) demostraron que se producen mejoras significativas en los pronósticos de inflación al entrenar sus modelos de *ML* con grandes conjuntos de datos. Así pues, sus modelos mostraron una mejora de 30% en términos de error cuadrático medio con respecto a modelos de referencia (Autorregresión y Caminata Aleatoria).

En el caso específico de México, Pedroza (2019) empleó redes neuronales como una herramienta innovadora para pronosticar la inflación. Sus investigaciones arrojaron resultados que respaldan la eficacia de estas redes como un modelo efectivo para obtener pronósticos estables sin la necesidad de llevar a cabo transformaciones complicadas en los datos. Así y todo, el autor señala que la calidad de los resultados dependerá en mayor medida del diseño del modelo.

La aportación principal de esta tesina a la literatura radica en un uso combinado de herramientas de aprendizaje de máquina y de un marco teórico especializado para la selección de variables que afectan la inflación en México a corto y largo plazo. Aunado a esto, se emplean modelos de Auto*ML* (Auto- Machine Learning) y se comparan con un modelo Long Short-Term Memory (*LSTM* de aquí en adelante), de este modo podremos comparar dos modelos dinámicos sin necesidad de asumir la forma funcional de los datos. Esta metodología permitirá evaluar si los modelos de Auto*ML*, que son más fáciles y rápidos de implementar, aportan eficiencia al proceso de ajustar el pronóstico.

El modelo *LSTM* fue elegido debido a su capacidad para capturar patrones y relaciones complejas en series de tiempo, especialmente en situaciones donde la dinámica de los datos es no lineal y puede estar influenciada por múltiples factores. Dado que la inflación está afectada por eventos económicos y políticas, los modelos *LSTM* resultan adecuados para captar los cambios en factores como tasas de interés, políticas fiscales o crisis económicas.

La principal característica por la que se eligieron a los modelos de Auto*ML* es su capacidad para realizar automáticamente tareas críticas en el ciclo de vida del modelado, como

la selección de características, la optimización de hiperparámetros y la evaluación del rendimiento del modelo. Esto significa que los usuarios pueden centrarse en la definición del problema y la preparación de datos, mientras que el Auto*ML* se encarga de gran parte del trabajo técnico.

La estructura de la tesina se divide en las siguientes secciones: revisión de literatura, descripción de los datos, metodología asociada a la implementación del modelo, resultados del análisis y conclusiones.

Capítulo 2

Revisión de literatura

El avance en el pronóstico de la inflación mediante métodos de *ML* ha crecido de manera exponencial en los últimos años. Desde tesis de maestría hasta artículos en revistas científicas y documentos de investigación de algunos bancos centrales. A este respecto, una de las investigaciones más mencionadas en la literatura es la de Medeiros et al (2019). Así pues, estos autores proponen que los métodos de *ML* motivados por una disponibilidad abundante en datos pueden ofrecer una mejora en el pronóstico de variables económicas tan importantes como la inflación. Ahondando en esto, los autores prueban que los modelos de *ML* superan a los modelos tradicionales en un 30%, en términos de reducción del error cuadrático medio (*MSE*).[1]

Cabe señalar que los primeros modelos de *ML* diseñados para pronosticar la inflación no lograron resultados significativos con relación a los modelos de referencia (autorregresivos principalmente). Ello se debe principalmente a que los modelos de *ML* utilizados para pronosticar procesos no lineales eran univariados y con datos escasos.[2] De este modo, los primeros modelos no podían captar el comportamiento complejo de los factores económicos asociados a estructuras multivariadas.

Ahora bien, dentro de estos primeros modelos de *ML* no-lineales univariados destaca el desarrollado por Nakamura (2004). En dicho artículo, el autor muestra cómo una red neuronal simple puede superar los resultados de un modelo autorregresivo univariado en un horizonte corto (1-2 trimestres). También concluye que, si bien las redes neuronales no sobrepasan los resultados de los modelos autorregresivos en todos los horizontes, habría que prestar mayor atención a técnicas de entrenamiento como el *early stopping* si se desea mejorar estos resultados en futuras investigaciones.

Esto se debe a que el *early stopping* es una técnica que detiene el entrenamiento de un modelo cuando su rendimiento en un conjunto de validación deja de mejorar, evitando el sobreajuste. Sus ventajas incluyen la prevención del sobreajuste, el ahorro de tiempo, recursos y la eficiencia

[1] Medeiros et al.(2019, p 2).
[2] Medeiros et al(2019, p. 3).

en la búsqueda de modelos óptimos. De hecho, esta técnica es de uso convencional en los modelos modernos de *ML*.

A pesar de que existe una amplia variedad de literatura con respecto al pronóstico de la inflación mediante modelos de *ML*, para el caso de México resulta escaza. Una interesante excepción es la del análisis realizado en Pedroza (2019), en el cual prueba un modelo de redes neuronales para el pronóstico de la inflación en México.

El autor concluye que las redes neuronales tienen la ventaja de proveer de un pronóstico consistente con relativa facilidad de implementación. Lo importante de las conclusiones de Pedroza es que, si bien no es necesario un complejo tratamiento de las series temporales, si lo es llevar a cabo un buen diagnóstico para poder elegir la mejor arquitectura del modelo.

Otro estudio para el caso de México es el de Chávez y Cortés (2013), los que implementaron múltiples redes neuronales, variando la cantidad de capas ocultas y el número de neuronas en cada capa. De forma tal que llegaron a un modelo de una sola capa oculta y 49 neuronas que les permitió superar al pronóstico del Banco de México.

Para otros países de América Latina, encontramos estudios del Banco Central de Chile, en el que Leal et al (2020) ponen aprueba diferentes modelos de *ML* para el pronóstico de la inflación en Chile. Es verdad que los autores del documento de trabajo no encuentran una ventaja significativa con respecto a los modelos de referencia, sin embargo, ellos mismos aceptan la limitación de su ejercicio debido a una falta de datos y de ajustes ex -post.

Por su parte, Alegre y Lozano (2022) aplican modelos de *Deep learning* (aprendizaje profundo) para el pronóstico del PIB en México, Colombia, Brasil, Argentina, Chile, y Perú. En lo particular, el PIB de México fue mejor pronosticado por uno de sus modelos. Este resultado alienta a pensar que los modelos de ML son herramientas que merecen ser revisadas para el propósito de pronosticar fenómenos económicos.

Capítulo 3

Descripción de los datos

Los datos utilizados para entrenar los modelos se dividen en variables objetivo y variables predictoras. Así, las variables objetivo son las series de tiempo de los niveles del Índice de Precios al Consumidor (general, subyacente y no subyacente) en México desde 1989 hasta 2022. Para elegir las variables predictoras, tomamos en cuenta el estudio de Esquivel y Razo (2003) acerca de las fuentes de inflación en México. En el mencionado trabajo, los autores concluyen que la dinámica de la inflación se ve afectada por varios factores como las variaciones en los salarios reales, la demanda de dinero y el tipo de cambio nominal. También influyen en ella componentes inerciales de la inflación interna, el incremento de los precios controlados por el gobierno mexicano y el diferencial entre las tasas de interés a corto plazo de México y Estados Unidos. En base a esto, las variables predictoras son divididas en las siguientes categorías:

Mercado laboral

- índice Nacional de Precios al Productor (*inpp, 2004-2023*), serie mensual, fuente: SIE Banxico
- Productividad media laboral por empleo de la manufactura (*pml, 2008-2023*), serie mensual, fuente: SIE Banxico
- Salario nominal (*sn, 2008-2023*), serie mensual, fuente: SIE Banxico
- Salario real (*sr, 2008-2023*), serie mensual, fuente: SIE Banxico

Mercado monetario

- Saldos monetarios reales M1 (*m1, 1989-2023*), serie mensual, fuente: SIE Banxico

Mercado cambiario y deuda del gobierno

- Tasa de rendimiento a 28 días CETES (*cetes28, 1989-2023*), serie mensual, fuente: SIE Banxico

- Tasa de rendimiento a 91 días CETES (*cetes91, 1989-2023*), serie mensual, fuente: SIE Banxico
- Tipo de cambio para solventar obligaciones denominadas en moneda extranjera (*e, 1989-2023*), serie mensual, fuente: SIE Banxico
- Diferencial entre el rendimiento de los CETES a 91 días y el rendimiento de los certificados del tesoro de Estados Unidos a 3 meses (*cetesdif, 1989-2023*), diferencial entre las series mensuales previamente mencionadas, fuente: SIE Banxico

Las series de cada variable fueron ajustadas a los requerimientos de cada modelo (imputación de datos faltantes, transformaciones, descomposición, etc.). Así mismo, la forma en la que se hacen las particiones de la base de datos para entrenar y probar el modelo se determina en base a cada modelo y el periodo a pronosticar.

Capítulo 4

Metodología

Basados en la metodología de Medeiros et al (2019) usamos un esquema de ventanas móviles de tamaño fijo. Así pues, ajustamos este esquema a cada modelo para pronosticar una ventana de 12 meses de inflación para los periodos de *2008* y *2022*.

Bajo este esquema, podemos representar la inflación de la siguiente forma:

$$\boldsymbol{\pi_{t+h=F_h(X_t)+\varepsilon_{t+h}}} \qquad \ldots\ldots(1)$$

en donde *t=1, 2, 3…T* representa los periodos mensuales, *h=1,2,3, … H* son los periodos mensuales hacia adelante y $\boldsymbol{X_t = (X_{1t}, X_{2t}, X_{3t}, \ldots X_{nt})}$ es el vector de *n* covariables que afectan a la inflación mediante la función $\boldsymbol{F_h}(*)$. Además, $\boldsymbol{\varepsilon_{t+h}}$ representa el error aleatorio de media cero.

De esta forma, la predicción de la inflación estará conformada de la siguiente manera:

$$\boldsymbol{\hat{\pi}_{t+h|t} = \hat{F}_{h,t-Rh+1:t}} \qquad \ldots\ldots(2)$$

Donde $\boldsymbol{\hat{F}_{h,t-Rh+1:t}}$ es la función objetivo estimada mediante los datos del periodo $t\text{-}R_h+1$ a *t* y R_h es el tamaño de la ventana.[3]

Los pronósticos de cada modelo son comparados mediante la métrica del error cuadrático medio (MSE por sus siglas en inglés).

4.1. Modelo de referencia: autorregresivo (AR)

La ecuación del pronóstico de inflación mediante el modelo AR es la siguiente:

[3] Medeiros et al(2019, p 4).

$$\widehat{\boldsymbol{\pi}}_{t+h|t} = \widehat{\boldsymbol{\beta}}_{0,h} + \widehat{\boldsymbol{\beta}}_{1,h}\boldsymbol{\pi}_t + \cdots + \widehat{\boldsymbol{\beta}}_{p,h}\boldsymbol{\pi}_{t-p+1} \qquad \ldots\ldots(3)$$

Para estimar el orden del modelo (p) utilizamos el criterio de información bayesiana (BIC) y el criterio de Akaike (AIC). En cuanto a la estimación de parámetros del modelo, utilizamos OLS.

Adicionalmente usamos como otra referencia a la media de los pronósticos (INPC general y subyacente) de la Encuesta sobre las Expectativas de los Especialistas en Economía del Sector Privado del Banco de México.

4.2. Modelos de aprendizaje de máquina

Dividiremos a los modelos de *ML* en dos grupos, el primero es un modelo con optimización de hiperparámetros manual y basado en literatura especializada para pronosticar la inflación. El segundo grupo es un conjunto de modelos contenidos en la estructura del aprendizaje automático, también conocido como Auto*ML*.

4.2.1. Optimización manual de hiperparámetros

En el ámbito del aprendizaje de máquina, los hiperparámetros son variables externas al modelo que, dependiendo de su ajuste, impactan directamente en su rendimiento. Estos parámetros no son aprendidos a partir de los datos, sino que deben ser configurados de forma manual. Un ejemplo común es cuántas neuronas y capas conforman una red neuronal.

Podemos comparar a los hiperparámetros con perillas que sirven para sintonizar la frecuencia de una radio. En este sentido, la optimización de hiperparámetros es la combinación de ajustes a las perillas que permiten sintonizar la frecuencia de una estación de radio de la forma más nítida y sin interferencia posible.

4.2.1.1. Modelo *LSTM*

Un modelo *LSTM* (*Long Short-Term Memory*), o Red Neuronal de Memoria a Largo Plazo, es un tipo de algoritmo de inteligencia artificial que se utiliza para realizar tareas que involucran secuencias de datos, como el procesamiento del lenguaje natural y la predicción de series

temporales. A diferencia de las redes neuronales tradicionales, las *LSTM* pueden recordar y utilizar información de eventos pasados en la secuencia, lo que las hace excelentes para comprender patrones y relaciones a lo largo del tiempo.

En el contexto del pronóstico de la inflación, un modelo *LSTM* se utiliza para analizar y predecir patrones en los datos históricos de la inflación a lo largo del tiempo. Dado que la inflación a menudo exhibe comportamientos secuenciales y dependencias a largo plazo, los modelos *LSTM* son especialmente adecuados para capturar estas relaciones y realizar pronósticos precisos.

Las celdas de memoria del *LSTM*, las estructuras más importantes del modelo permiten retener información relevante de periodos anteriores y olvidar información menos relevante. Esto significa que el modelo puede aprender patrones a largo plazo en los datos de inflación, como estacionalidad, tendencias y ciclos económicos. Al tener en cuenta estas dependencias a largo plazo, el modelo *LSTM* puede hacer predicciones más precisas y capturar cambios abruptos o fluctuaciones en la inflación.

Al entrenar el modelo *LSTM* con datos históricos de inflación y sus correspondientes valores futuros, el modelo aprende a ajustar los pesos de las conexiones neuronales para minimizar el error de predicción. Una vez entrenado, el modelo puede utilizarse para realizar pronósticos futuros de inflación en función de nuevos datos de entrada.

Ahondando en el funcionamiento interno del modelo, podemos identificar las siguientes etapas:

Entrada de secuencia: El modelo *LSTM* toma una secuencia de datos de entrada, en este caso, los valores históricos de inflación a lo largo del tiempo (x_t). Cada valor de inflación se considera un punto en la secuencia temporal.

Celda de memoria: La celda de memoria es el componente principal de un modelo *LSTM* dado que funciona como un mecanismo de retención y control de información. Esta celda tiene la capacidad de recibir datos de entrada y decidir si esos datos deben ser guardados o descartados en función de su relevancia histórica. Así pues, una celda de memoria es estructurada en un estado de memoria inicial (c_t) y un estado oculto (h_t).

Podemos comparar el estado de memoria inicial (c_t) de una celda *LSTM* con un pizarrón en blanco que está listo para recibir información. En cuanto al estado oculto (h_t), podríamos imaginarlo como un asistente silencioso que observa lo que sucede en el pizarrón (el estado de

memoria). A medida que se van escribiendo y borrando cosas en el pizarrón, este asistente lleva un registro y mantiene una representación interna de lo que ha sucedido, incluso de cosas que ya han sido borradas. Así, el estado oculto actúa como una especie de "memoria a corto plazo" que ayuda a la red a entender mejor la información actual y pasada en el proceso de pronosticar la inflación.

Puertas: En cada paso de tiempo, el modelo *LSTM* utiliza diferentes puertas para controlar el flujo de información dentro de la celda de memoria. Estas puertas son redes neuronales que operan en los datos de entrada y en el estado oculto anterior. Según Sotaquirá (2023), estas puertas funcionan como válvulas; cuando están totalmente abiertas (valores cercanos a 1) la información pasa libre y cuando están totalmente cerradas (valores iguales a cero) la bloquean.

Puerta de entrada (*input gate*): Determina qué información nueva debe ser almacenada en la celda de memoria. Esta puerta procesa la entrada actual (x_t) y el estado oculto anterior (h_{t-1}) para calcular una "activación de entrada" que pondera la información entrante. Esta puerta genera el vector $\boldsymbol{i_t} = \boldsymbol{\sigma}(\boldsymbol{W_i x_t} + \boldsymbol{U_i h_{t-1}} + \boldsymbol{b_i})$, donde $\boldsymbol{\sigma}(*)$ es una función sigmoidal, de tal forma que los parámetros $(\boldsymbol{W_i}, \boldsymbol{U_i}, \boldsymbol{b_i})$ son estimados durante el entrenamiento.

Puerta de olvido (*forget gate*): Regula qué información anterior se debe olvidar en la celda de memoria. Esta puerta considera la entrada actual y el estado oculto anterior para calcular una "activación de olvido", de esta manera se decide qué información almacenada previamente debe ser descartada. Dicha puerta genera el vector $\boldsymbol{f_t} = \boldsymbol{\sigma}(\boldsymbol{W_f x_t} + \boldsymbol{U_f h_{t-1}} + \boldsymbol{b_f})$ donde $\boldsymbol{\sigma}(*)$ es una función sigmoidal y los parámetros $(\boldsymbol{W_f}, \boldsymbol{U_f}, \boldsymbol{b_f})$ son estimados durante el entrenamiento.

Puerta de salida (*output gate*): Determina qué información de la celda de memoria debe ser utilizada como salida. Utilizando la entrada actual (x_t) y el estado oculto anterior (h_{t-1}), esta puerta genera una "activación de salida" que selecciona qué información de la celda de memoria se debe utilizar para la predicción en ese paso de tiempo. La puerta de salida genera el vector $\boldsymbol{o_t} = \boldsymbol{\sigma}(\boldsymbol{W_o x_t} + \boldsymbol{U_o h_{t-1}} + \boldsymbol{b_o})$ donde $\boldsymbol{\sigma}(*)$ es una función sigmoidal y los parámetros $(\boldsymbol{W_0}, \boldsymbol{U_o}, \boldsymbol{b_o})$ son estimados durante el entrenamiento.

Actualización de la celda de memoria: La celda de memoria actualiza su estado en función de las activaciones de entrada y olvido calculadas en las puertas anteriores. La celda de

memoria combina la información anterior con la nueva información ponderada para obtener un estado de celda de memoria actualizado. Para ponderar la información de entrada es necesario crear otro vector activado por una función tangente hiperbólica. Este vector es de la forma: $\tilde{\boldsymbol{c}}_{\boldsymbol{t}} = \boldsymbol{tanh}(\boldsymbol{W}_{\boldsymbol{c}}\boldsymbol{x}_{\boldsymbol{t}} + \boldsymbol{U}_{\boldsymbol{c}}\boldsymbol{h}_{\boldsymbol{t-1}} + \boldsymbol{b}_{\boldsymbol{c}})$. La celda de estado actualizado está determinada por la siguiente ecuación:

$\boldsymbol{c}(\boldsymbol{t}) = [\boldsymbol{f}_{\boldsymbol{t}} \times \boldsymbol{c}(\boldsymbol{t}-\boldsymbol{1})] + [\boldsymbol{i}_{\boldsymbol{t}} \times \tilde{\boldsymbol{c}}_{\boldsymbol{t}}]$, donde × denota el producto punto de los vectores señalados en la ecuación.

Cálculo del estado oculto: El estado oculto se calcula utilizando la activación de salida de la puerta de salida (o_t) y el estado de celda de memoria actualizado ($c(t)$). El estado oculto captura información relevante para la predicción en ese paso de tiempo y también se utiliza como entrada para el siguiente paso de tiempo. De esta forma, el estado oculto queda representado como: $\boldsymbol{h}_{\boldsymbol{t}} = \boldsymbol{o}_{\boldsymbol{t}} \times \boldsymbol{tanh}[\boldsymbol{c}(\boldsymbol{t})]$.

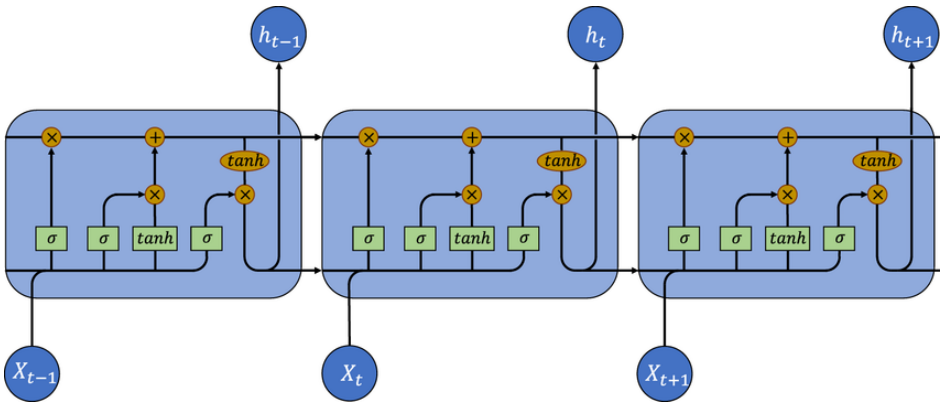
Predicción: Finalmente, el estado oculto actualizado se utiliza para realizar la predicción de inflación en el paso de tiempo actual. De tal forma que la predicción mediante este modelo queda representada con la siguiente expresión:

$$\hat{\boldsymbol{\pi}}_{\boldsymbol{t+h|t}} = \boldsymbol{W}_{\boldsymbol{y}}\boldsymbol{h}_{\boldsymbol{t}} + \boldsymbol{b}_{\boldsymbol{y}} \qquad \text{…. (4)}$$

La Figura 1 representa la estructura secuencial de las etapas previamente mencionadas, donde la línea horizontal que recorre el esquema por arriba corresponde a los estados de la celda ***c(t)***. Mientras que la línea horizontal inferior corresponde al estado oculto $\boldsymbol{h}_{\boldsymbol{t}}$, de modo que ambos estados alimentan a la celda del siguiente periodo junto con las entradas de secuencia $\boldsymbol{x}_{\boldsymbol{t}}$. Además, cada línea recorre un camino de mecanismos y operaciones que dan lugar a la actualización de la celda y estado oculto en cada periodo.

El primer cuadro con el símbolo $\boldsymbol{\sigma}$, de izquierda a derecha, representa a la puerta de olvido $\boldsymbol{f}_{\boldsymbol{t}}$, el siguiente bloque activado mediante $\boldsymbol{\sigma}$ representa la puerta de entrada $\boldsymbol{i}_{\boldsymbol{t}}$ y el último bloque $\boldsymbol{\sigma}$ la puerta de salida $\boldsymbol{o}_{\boldsymbol{t}}$. Cabe destacar que × denota el producto punto de los vectores señalados en el esquema, $\boldsymbol{\sigma}$ representa la función de activación sigmoidal y ***tanh*** representa la función de activación tangente hiperbólica.

Figura 1: Esquema del modelo *LSTM*



Fuente: R. Peirano et al (2021). El esquema representa la estructura y funcionamiento interno del modelo *LSTM* en cada periodo de tiempo.

4.2.2. Optimización automática de hiperparámetros: Auto*ML*

El Auto*ML* busca automatizar tareas clave del flujo de trabajo de *ML*, como la selección de atributos, la ingeniería de atributos, la selección del modelo, la optimización de hiperparámetros y la evaluación del rendimiento del modelo. Según Zhen Xu et al.(2021) el Auto*ML* surge ante la necesidad de reducir la intervención humana en el proceso de crear y ejecutar modelos de *ML* ante tareas específicas. En base a lo anterior, podemos decir que el objetivo principal del Auto*ML* es reducir la barrera de entrada y hacer que el aprendizaje automático sea más accesible incluso para aquellos que no tienen una experiencia profunda en ciencia de datos.

En el contexto de Auto*ML* existen diferentes paqueterías y herramientas, pero elegimos AutoGluon (2020) por sus buenos resultados en pronosticar series de tiempo.[4] En este sentido, AutoGluon brinda un catálogo de modelos que son entrenados con unas pocas líneas de código. A continuación, se presenta el catálogo y una breve descripción de cada modelo en la Tabla 1.

[4] Zhen Xu et al.(2021, p 2).

Tabla 1: Catálogo de AutoGluon para series de tiempo

Modelo	**Descripción**
NaiveModel	Modelo de referencia que establece la previsión igual al último valor observado
SeasonalNaiveModel	Modelo de referencia que establece la previsión igual al último valor observado de la misma temporada
ARIMAModel	Autoregressive Integrated Moving Average (ARIMA) model.
ETSModel	Exponential smoothing con tendencia y estacionalidad
ThetaModel	El modelo Theta de Assimakopoulos y Nikolopoulos (2000)
AutoETSModel	Selecciona automáticamente el mejor modelo ETS (Error, Tendencia, Estacionalidad) utilizando un criterio de información
AutoARIMAModel	Selecciona automáticamente los mejores parámetros del modelo (p,d,q,P,D,Q) utilizando un criterio de información
DynamicOptimizedThetaModel	Modelo Theta optimizado de Fiorucci et al. (2016)
AutoGluonTabularModel	Pronóstico de series de tiempo mediante autogluon.tabular.TabularPredictor
DeepARModel	El modelo consta de un codificador *LSTM* y un decodificador que emite la distribución del siguiente valor objetivo
SimpleFeedForwardModel	El modelo consiste en un perceptrón multicapa (*ML*P) que predice la distribución de todo el valor objetivo en el horizonte de previsión

TemporalFusionTransformerModel	El modelo combina un codificador *LSTM*, un decodificador transformer y predice directamente los cuantiles de los valores objetivo-futuros

Fuente: AutoGluon – Model Zoo.

Capítulo 5

Resultados

Para llevar a cabo los pronósticos de la inflación, se implementó una estrategia de partición de las series de tiempo. En el proceso de entrenamiento del modelo *LSTM*, se utilizó exclusivamente la serie del INPC (general, subyacente y no subyacente según sea el caso) desde 1969 hasta 2012, mientras que el conjunto de prueba corresponde el periodo 2013-2022. Es relevante señalar que entrenamos un modelo por cada componente del INPC.

Por otro lado, para los modelos de Auto*ML* se empleó el INPC como variable objetivo y múltiples variables predictoras con el periodo de entrenamiento 1989-2021 y el periodo de prueba 2022. Las variables predictoras abarcan diferentes aspectos del mercado laboral, monetario, cambiario y deuda del gobierno. Estas series incluyeron el Índice de Precios al Productor (***ipp***), productividad media laboral (***p*ml**), salario nominal (***sn***), salario real (***sr***), saldos monetarios reales (***m1***), tasa de rendimiento a 28 días CETES (***cetes28***), tasa de rendimiento a 91 días CETES (***cetes91***), tipo de cambio para solventar obligaciones denominadas en moneda extranjera (***e***) y el diferencial entre el rendimiento de los CETES a 91 días y el rendimiento de los Certificados del Tesoro de Estados Unidos a 3 meses (***cetesdif***).

Cabe destacar que esta partición del conjunto de entrenamiento y validación fue elegida en base a la evaluación del rendimiento de cada modelo en diferentes particiones. Así, el modelo AR fue entrenado exclusivamente con la serie del INPC (en sus diferentes componentes) con el periodo de entrenamiento 1989-2021 y el periodo de prueba 2022. Además, determinamos el orden del modelo mediante el criterio de información bayesiana (BIC) y el criterio de Akaike (AIC).

Una vez que los modelos fueron entrenados para cada serie del INPC (general, subyacente y no subyacente) y de haber verificado que ninguno presentara sobreajuste, se realizaron los pronósticos anuales para el periodo 2022 y 2008. En la siguiente sección mostramos los resultados para cada periodo desglosado en INPC general, subyacente y no subyacente.

*Todos los modelos fueron programados en Python y el código es de libre acceso mediante el enlace: https://github.com/eriknow95/*LSTM*_AutoML_Forecast_INPC_MX.git

5.1. Predicciones para 2022

5.1.1. INPC General

En el análisis de las predicciones para el INPC General en 2022, se observó un rendimiento sobresaliente por parte de los modelos de *ML* en comparación con el modelo de referencia AR (autorregresivo) y la media de los pronósticos de la Encuesta sobre las Expectativas de los Especialistas en Economía del Sector Privado del Banco de México (Banxico). A continuación, se presenta una tabla que muestra el Error Cuadrático Medio (MSE) de los distintos modelos:

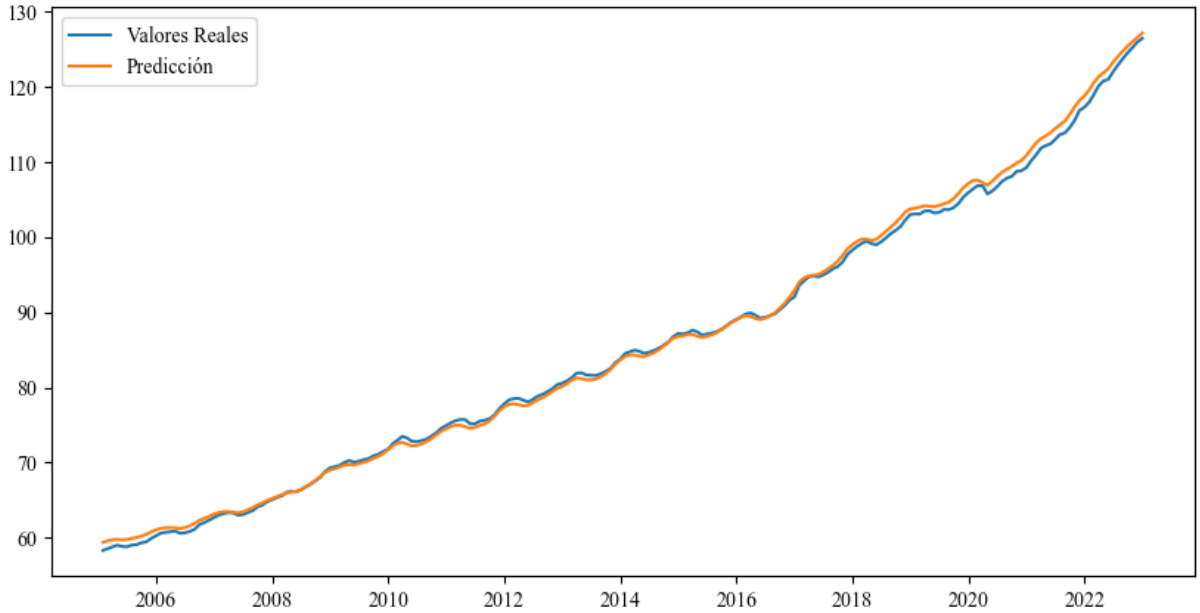
Tabla 2: MSE para el pronóstico del INPC General (2022)

	Modelo	MSE
1	*LSTM*	1.3377
2	Theta-*AML*	1.5990
3	TemporalFusionTransformer-*AML*	2.6494
4	ETS-*AML*	7.2985
5	ARIMA-*AML*	7.4595
6	AR	13.1631
7	Banxico	15.1509
8	WeightedEnsemble-*AML*	23.3561
9	SeasonalNaive-*AML*	27.2405
10	Naive-*AML*	27.2405

Fuente: elaboración propia con datos del SIE Banxico. Tanto los modelos de *ML* como el de referencia superaron al pronóstico medio de la encuesta de Banxico.

Como se puede observar en la Tabla 1, el modelo *LSTM* destacó con el *MSE* más bajo, seguido por varios modelos de *Auto Machine Learning* (Theta-*AML*, *Temporal Fusion Transformer-AML*, ETS-*AML* y ARIMA-*AML*), que también superaron ampliamente al AR y a la media de Banxico en términos de MSE. Aunado a esto, podemos comprobar en la Figura 2 que el modelo *LSTM* pronostica de manera efectiva el comportamiento de la serie en horizontes temporales más amplios (en este caso de 17 años).

Figura 2: Predicción del INPC General en el periodo 2005-2022 mediante el modelo *LSTM*



Fuente: elaboración propia con datos del SIE Banxico. Podemos observar que el modelo es capaz de pronosticar el comportamiento de la serie inclusive en periodos de saltos abruptos como lo fue 2020-2021.

En términos de reducción del MSE, el modelo *LSTM* representó una mejora del 89% con respecto al modelo AR, 91% con respecto a la encuesta de Banxico y 16% con respecto al Auto*ML*. Este resultado sugiere que los modelos *LSTM* mejoran significativamente la precisión del pronóstico. Además, el marco del Auto*ML* presenta una alternativa eficiente en términos de recursos, tiempo y resultados.

5.1.2. INPC Subyacente

Al analizar las predicciones para el INPC Subyacente en 2022, se observó un predominio de los modelos de *Machine Learning*, que ocuparon los seis primeros lugares en términos de MSE. A continuación, se presenta una tabla con los resultados:

Tabla 3: MSE para el pronóstico del INPC Subyacente (2022)

	Modelo	MSE
1	*LSTM*	0.1422
2	ETS-*AML*	0.1535
3	ARIMA-*AML*	0.1546
4	DeepAR-*AML*	0.4015
5	Theta-*AML*	7.8858
6	TemporalFusionTransformer-*AML*	12.880001
7	Banxico	13.3659
8	AR	18.2689578
9	SeasonalNaive-*AML*	35.7262
10	Naive-*AML*	35.7262377

Fuente: elaboración propia con datos del SIE Banxico. Los modelos de *ML* superaron al pronóstico medio de la encuesta de Banxico.

Según el informe anual de Banxico (2023) para el año 2022, los choques del efecto rezagado de la pandemia y el conflicto de Rusia-Ucrania aumentaron las presiones inflacionarias particularmente en los alimentos. Este hecho se tradujo en un mayor grado de incertidumbre y en que los modelos de *ML* pudieron pronosticar con menor margen de error el componente subyacente de la inflación. Así lo podemos constatar en la Tabla 3, donde los modelos *LSTM* y ETS-*AML* redujeron el MSE en 98% con respecto a la encuesta de Banxico.

5.1.3. INPC No Subyacente

Para el INPC No Subyacente en 2022, se presenta la Tabla 4 que muestra los resultados de MSE:

Tabla 4: MSE para el pronóstico del INPC No Subyacente (2022)

	Modelo	MSE
1	*LSTM*	1.2563
2	TemporalFusionTransformer-*AML*	8.1407
3	SeasonalNaive-*AML*	8.9302
4	Naive-*AML*	8.9302
5	AR	9.1993
6	Theta-*AML*	13.5202704
7	WeightedEnsemble-*AML*	19.7318
8	DeepAR-*AML*	59.7083727
9	ARIMA-*AML*	97.1314
10	ETS-*AML*	98.157335

Fuente: elaboración propia con datos del SIE Banxico. Los modelos de *ML* tuvieron mayor éxito al pronosticar la dinámica volátil del INPC no subyacente.

Como se puede observar en la Tabla 4, el modelo *LSTM* obtuvo el MSE más bajo, lo cual representó una reducción del 86% con respecto al margen de error del modelo AR. Así y todo, el modelo Temporal Fusion Transformer-*AML* sólo representó una reducción del 11% con respecto al MSE del modelo AR. Las figuras contenidas en el apéndice de esta tesina ilustran los pronósticos de estos modelos.

5.2. Predicciones para 2008

Es necesario aclarar que para la siguiente sección de resultados tomamos el modelo de Auto Machine Learning que mostró el mejor rendimiento en la fase de entrenamiento y prueba a modo de resumen. De este modo, el mejor modelo de Auto Machine Learning queda representado por la palabra clave: Auto*ML*. El nombre del modelo específico y los pronósticos asociados se encuentran en la sección de apéndice de esta tesina.

5.2.1. INPC General

En el análisis de las predicciones para el INPC General en 2008, se observó que el modelo *LSTM* ocupó el primer lugar en términos de MSE, seguida por la media de la encuesta de Banxico y el modelo AR. El modelo de Auto Machine Learning (*AML*) se ubicó en último lugar en términos de precisión. Se puede consultar la Tabla 5 a continuación para una visión detallada de estos resultados:

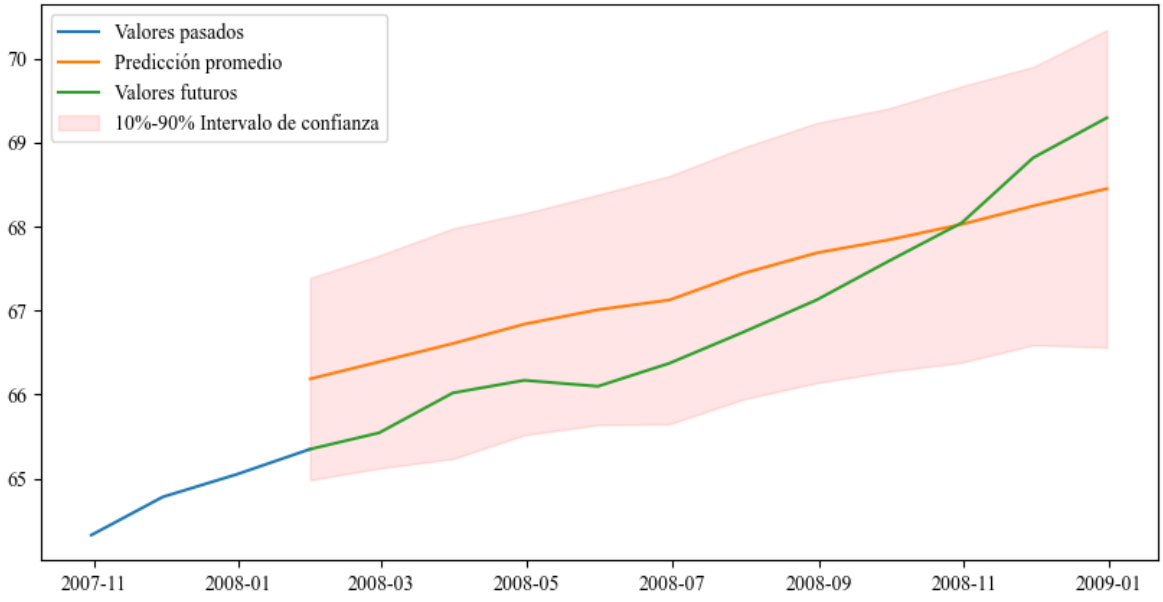
Tabla 5: MSE para el pronóstico del INPC General (2008)

	Modelo	MSE
1	*LSTM*	0.0159
2	Banxico	0.0321
3	AR	0.1964
4	Auto*ML*	0.4597

Fuente: elaboración propia con datos del SIE Banxico.

El rendimiento del modelo *LSTM* se traduce en una reducción del 50% y 92% en términos del MSE de la encuesta de Banxico y el modelo AR, respectivamente. Sin embargo, el modelo de Auto*ML* no superó a ningún modelo. De cualquier forma, la Figura 3 muestra cómo el pronóstico del modelo de Auto*ML* contiene a los valores reales del INPC en sus intervalos de confianza, por lo que todavía no podemos rechazar la efectividad de este modelo.

Figura 3: Predicción del INPC General en el periodo 2008 mediante el modelo Ensamble-Auto*ML*



Fuente: elaboración propia con datos del SIE Banxico.

5.2.2. INPC Subyacente

Para el INPC Subyacente en 2008, se presenta la Tabla 6 con los resultados de MSE de cada modelo:

Tabla 6: MSE para el pronóstico del INPC Subyacente (2008)

	Modelo	MSE
1	*LSTM*	0.0111
2	Banxico	0.0116
3	Auto*ML*	0.0674
4	AR	0.0818

Fuente: elaboración propia con datos del SIE Banxico.

El modelo *LSTM* mejoró la precisión del pronóstico en 4.3% con respecto al MSE de la encuesta de Banxico y el modelo Auto*ML* superó al modelo de referencia AR. Esta mejora

marginal puede ser atribuida a que, según el informe anual del Banco de México (2009) para el año 2008, las presiones inflacionarias se vieron reducidas debido a la baja actividad económica. Además, el aumento de inflación a finales de ese año se vio especialmente reflejado en los precios de las materias primas, por lo que el componente subyacente de la inflación no presentó cambios abruptos que los modelos convencionales no pudieran detectar.

5.2.3. INPC No Subyacente

Finalmente, para el INPC No Subyacente en 2008, se presenta la Tabla 6, en la que se muestran los resultados de MSE:

Tabla 7: MSE para el pronóstico del INPC No Subyacente (2008)

	Modelo	MSE
1	*LSTM*	0.7582
2	Auto*ML*	0.9845
3	AR	1.3888

Fuente: elaboración propia con datos del SIE Banxico.

Como era de esperarse para el pronóstico del componente más volátil de la inflación, los modelos de *ML* superaron a los pronósticos del modelo de referencia AR. En términos de reducción del MSE, los modelos *LSTM* y Auto*ML* superaron al modelo AR en 45% y 29%, respectivamente.

En resumen, los resultados presentados en estas tablas indican que, en el período de 2022, los modelos de *ML* superaron por un amplio margen al modelo de referencia AR y a la media de los pronósticos recabados por la encuesta de Banxico en términos de precisión. Sin embargo, para el período de 2008, los modelos de Auto*ML* no reflejaron una mejora consistente como sí lo fue el *LSTM*.

Capítulo 6

Conclusión

El objetivo central de esta tesina fue examinar las ventajas que brindan los modelos de aprendizaje de máquina en el pronóstico de la inflación mexicana durante periodos de incertidumbre económica, centrándonos específicamente en la crisis financiera de 2007-2008 y el periodo de postpandemia de 2022. Para lograr esto, comparamos diferentes modelos de *ML* con un modelo autorregresivo (AR) y con los resultados de la Encuesta sobre las Expectativas de los Especialistas en Economía del Sector Privado del Banco de México.

Los resultados obtenidos en esta investigación respaldan la hipótesis planteada. Durante el período de análisis de 2022, los modelos de *ML* demostraron un rendimiento sobresaliente al superar ampliamente al modelo referencia AR y a la media de los pronósticos recabados por la encuesta de Banxico. En particular, el modelo *LSTM* se destacó al mostrar los MSE más bajos en la predicción de la inflación general, subyacente y no subyacente.

El rendimiento de los modelos *LSTM* para el periodo 2022 representan una mejora del 91% en la precisión del pronóstico del INPC general con respecto a la encuesta del Banco de México. También, una mejora del 98% a la precisión del pronóstico del componente subyacente de la inflación (con respecto a la encuesta de Banxico) y una mejora del 86% en el pronóstico del INPC no subyacente (con respecto al modelo AR).

No obstante, al considerar el período de 2008 se observó que, si bien los modelos *LSTM* superaron de forma consistente a los modelos de referencia, no fuel el caso de los modelos de Auto*ML* que sólo superaron al modelo AR en pronosticar el componente subyacente y no subyacente de la inflación. Este resultado sugiere que la precisión de los modelos de Auto*ML* puede variar según las condiciones económicas y financieras particulares de cada período de incertidumbre.

En este respecto, el informe anual del Banco de México (2009) señala que las presiones inflacionarias disminuyeron en 2008 debido a la baja actividad económica. El documento también afirma que el aumento de la inflación al final del año se debió en gran medida al aumento en los precios de las materias primas, lo cual reforzó el comportamiento no lineal de

la inflación no subyacente. En conclusión, 2008 no presentó un panorama tan volátil en precios como el periodo de 2022 después del inicio de la guerra de Ucrania.

Los resultados obtenidos para 2008 son consistentes con la hipótesis planteada y con el panorama específico de ese periodo ya que el modelo *LSTM* mejoró el pronóstico de la encuesta del Banco de México para el INPC general en 50%. Además, mejoró marginalmente (4.3%) la precisión del pronóstico para el componente subyacente debido a que no hubo movimientos abruptos en la serie que los modelos convencionales no pudieran captar.

Por último, todos los modelos de aprendizaje de máquina (*LSTM* y Auto*ML*) superaron por un amplio margen (45%) al modelo de referencia AR en la precisión del pronóstico del INPC no subyacente, componente que se vio afectado por el incremento en los precios de las materias primas ese año. Aún si los modelos de Auto*ML* no pudieron superar de forma consistente a los modelos de referencia en este periodo, debemos tomarlos en cuenta ya que obtuvieron buenos resultados al pronosticar el componente más volátil de la inflación y su uso puede ahorrarnos tiempo y recursos.

En cuanto a la eficiencia, tanto el modelo *LSTM* como el modelo *AR* requieren un tiempo y recursos computacionales similares para su entrenamiento e implementación. En ambos casos, la especificación de variables, la preparación de datos y la codificación para ejecutar los modelos son las tareas que consumen más tiempo, con un total de 2 semanas para este ejercicio específico. Además, el tiempo de ejecución del código fue de solo unos minutos.

En comparación, la paquetería de Auto*ML* logra realizar todo el proceso, incluyendo el tratamiento de datos, el entrenamiento, la prueba y la ejecución del código, en aproximadamente 20 minutos. De esta manera, la paquetería de Auto*ML* representa un ahorro significativo de tiempo en comparación con los modelos mencionados anteriormente.

La principal limitación de este estudio es la falta de un mayor número de datos históricos sobre la inflación en México. Aunque tenemos disponible información mensual de la inflación general de México desde 1969, esto sólo representa 636 datos de entrada para entrenar al modelo, lo cual resulta poco en comparación con los miles de millones en el conjunto de entrenamiento de los modelos más poderosos de la Inteligencia Artificial. Para tratar de remediar esta situación, fueron usadas diferentes variables predictoras que le dieran una

dimensión mayor de aprendizaje al modelo de Auto*ML*, sin embargo, estas variables eran aún más limitadas en extensión.

Otra limitación de esta tesina es la falta de interpretabilidad de los modelos de *ML*, de modo que podemos determinar la efectividad de estos mediante sus resultados, pero no podemos describir en qué medida una variable o algún rezago afecta en mayor medida a los resultados. Es verdad que hoy en día existen meta-algoritmos que permiten interpretar a los modelos de *ML*, como el algoritmo de gradientes integrados que se utiliza para entender cómo las características de entrada contribuyen a las predicciones realizadas por el modelo, pero su implementación escapa a las habilidades de programación del autor y a la extensión de esta tesina.

De cualquier forma, los hallazgos de esta tesina respaldan la utilidad y la eficacia de los modelos de aprendizaje de máquina como herramientas valiosas para el pronóstico de la inflación en contextos de incertidumbre económica, como el que caracterizó al año 2022. Estos modelos tienen la capacidad de captar relaciones no-lineales en los datos y adaptarse a las complejas dinámicas económicas, lo que los hace más adecuados para enfrentar desafíos pronósticos en entornos económicos volátiles. Sin embargo, es importante destacar que la elección del modelo adecuado en el marco del *Auto Machine Learning* puede depender de las condiciones específicas del momento y del tipo de variable que se esté pronosticando.

Este estudio contribuye al análisis empírico de la economía mexicana y a las metodologías de predicción económica al proporcionar evidencia de la efectividad de los modelos de *ML* en el pronóstico de la inflación en contextos de incertidumbre. Con estas evidencias, los analistas pueden enriquecer el conjunto de herramientas disponibles para realizar sus predicciones. Además, subraya la importancia de considerar el contexto económico y la naturaleza de los datos al seleccionar el enfoque de modelado más adecuado para el pronóstico de la inflación en tiempos de incertidumbre económica.

Capítulo 7

Bibliografía

Alegre Ibáñez, Augusto. Lozano Aparicio, José Martin. *"Aplicación de métodos de Deep Learning en series de tiempo para el pronóstico de la situación macroeconómica en América Latina"*. Interfases, No. 15. Universidad de Lima (2022). pp. 127-128.

Assimakopoulos, Vassilis. Nikolopoulos, Konstantinos. *"The theta model: a decomposition approach to forecasting"*. International journal of forecasting, No. 16, Vol. 4 (2000). pp. 521-530.

AutoGluon, *"Forecasting Time Series – Model Zoo"*, Model Zoo, https://auto.gluon.ai/stable/tutorials/timeseries/forecasting-model-zoo.ht*ML* , consultado el 25 de mayo de 2023.

Banco de México. *"Compilación de Informes Trimestrales Correspondientes al año 2022"*. Informe Anual y compilación de informes trimestrales (2023). pp 1-340.

Banco de México. *"Resumen Informe Anual 2008"*. Informe Anual y compilación de informes trimestrales (2009). pp 1-16.

Chávez Hurtado, José Luis. Cortés Fragoso, José Héctor. *"Forecasting Mexican inflation using neural networks"*, CONIELECOMP (2013). pp. 32-35.

Esquivel, Gerardo. Razo, Raúl. *"Fuentes de inflación en México, 1989-2000: un análisis multicausal de corrección de errores"*. Estudios Económicos, Vol. 18, No. 2 (36). Colegio de México (2003), pp. 182-215.

Fiorucci, Jose. Pellegrini, Tiago. Louzada, Francisco. Petropoulos, Fotios. Koehler, Anne. *"Models for optimising the theta method and their relationship to state space models"*. International journal of forecasting No. 32, Vol. 4 (2016). pp.1151-1161.

Leal, Felipe. Molina, Carlos. Zilberman, Eduardo. *"Proyección de la inflación en Chile con modelos de ML"*. Banco Central de Chile (2020), documentos de trabajo, No. 860. pp 1-17.

Medeiros, Marcelo. Vasconcelos, Gabriel. Veiga, Alvaro. "*Forecasting Inflation In Data Rich Environments*". Journal of Business and Economic Statistics (2019). pp 2-40.

Nakamura, Emi. "*Inflation forecasting using a neural network*". Department of economics, Harvard University (2004). pp 374-377.

Pedroza Robles, Héctor. *"Uso de redes neuronales para el pronóstico de la inflación en México"*, Almanaque Económico, No.1.Colegio de México (2019). pp 63-67.

Peirano, R., Kristjanpoller, W. & Minutolo, M.C. *"Forecasting inflation in Latin American countries using a SARIMA–LSTM combination"*. *Soft Computing* **25**, 10851–10862 (2021). https://doi.org/10.1007/s00500-021-06016-5.

Sotaquirá, Miguel. *"¿Qué son las Redes LSTM?"*. Codificando Bits, consultado el 1 de enero de 2023.

Xu, Zhen. Tu, Wei-Wei. Guyon, Isabelle. *"AutoML meets Time Series Regression"*. ArXiv. Cornell University (2021), https://arxiv.org/abs/2107.13186 , consultado el 22 de marzo de 2023, pp.2.

Capítulo 8

Apéndice

8.1. Pronósticos del modelo *LSTM*

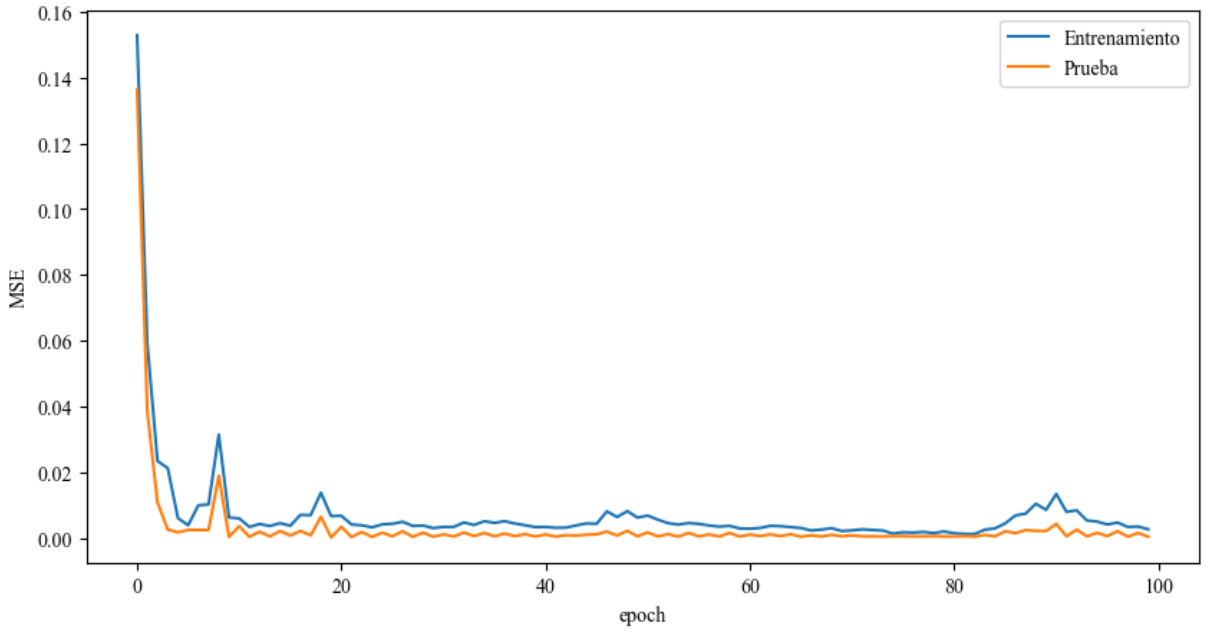
8.1.1. Predicciones 2022

8.1.1.1. INPC General

La Figura 4 ilustra la calibración del modelo *LSTM* en cien épocas para el conjunto de entrenamiento y el conjunto de prueba. Cabe destacar que el modelo no presenta sobreajuste ya que el error en el conjunto de prueba fue menor que en el conjunto de entrenamiento después de las primeras épocas. Una época representa un ciclo de entrenamiento con los datos de entrenamiento y de prueba.

De este modo, el modelo aprendió los patrones de la serie y realiza pronósticos con un bajo MSE ante nuevos datos.

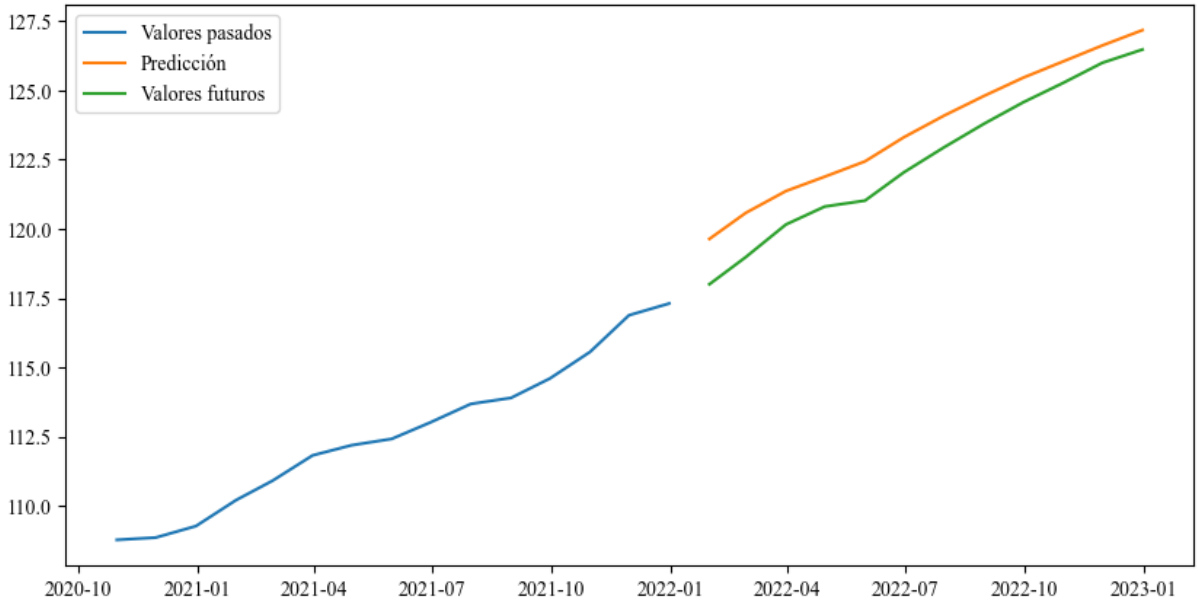
Figura 4: Calibración del modelo *LSTM* para pronosticar el INPC General



Fuente: elaboración propia con datos del SIE Banxico.

Podemos apreciar la predicción del INPC general para el periodo 2022 en la Figura 5, la cual proyecta una expectativa al alza en el nivel de precios para dicho periodo. Esto corresponde con el comportamiento de los precios inclusive en los primeros meses de 2023 debido a los efectos rezagados de la pandemia y el conflicto entre Rusia y Ucrania.

Figura 5: Predicción del INPC General en el periodo 2022 mediante el modelo *LSTM*

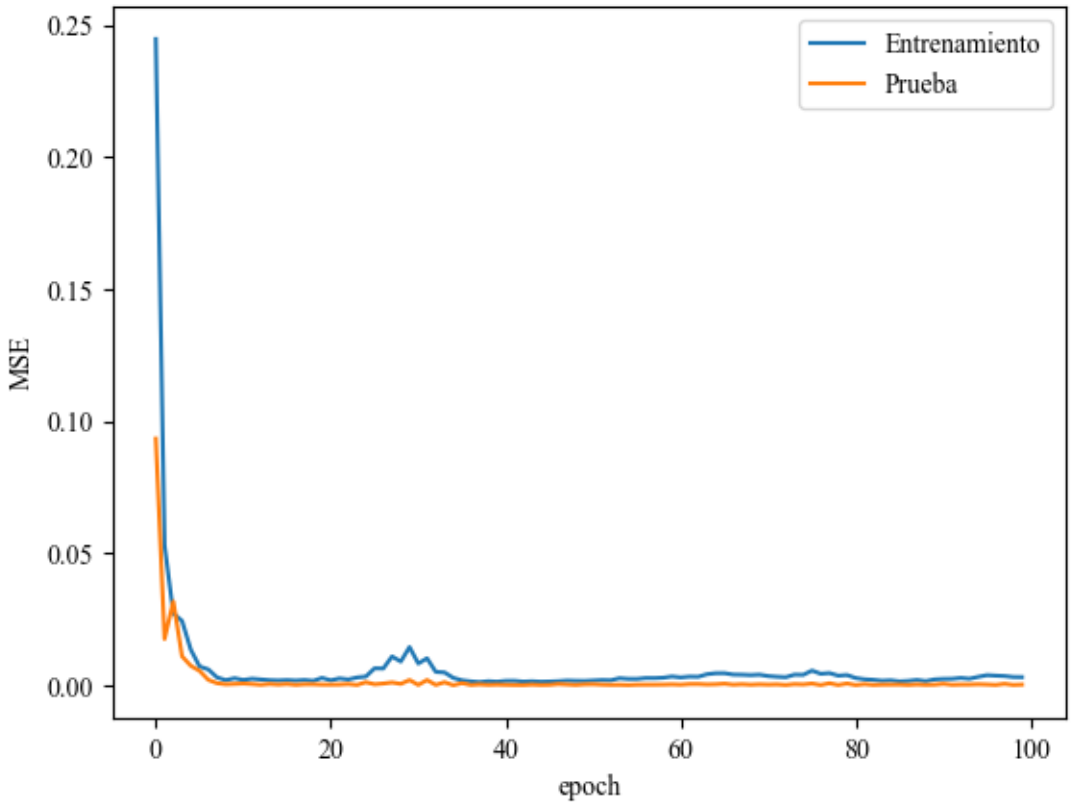


Fuente: elaboración propia con datos del SIE Banxico.

8.1.1.2. INPC Subyacente

La Figura 6 muestra la calibración del modelo *LSTM* para pronosticar el INPC subyacente. En esta figura observamos un bajo MSE para el conjunto de entrenamiento y prueba durante la calibración, lo cual sugiere que el modelo no presenta sobreajuste. Así lo podemos comprobar en la Figura 3 de la sección de resultados, en la que se muestra cómo el modelo pronostica fielmente el comportamiento de la serie para un horizonte de 17 años.

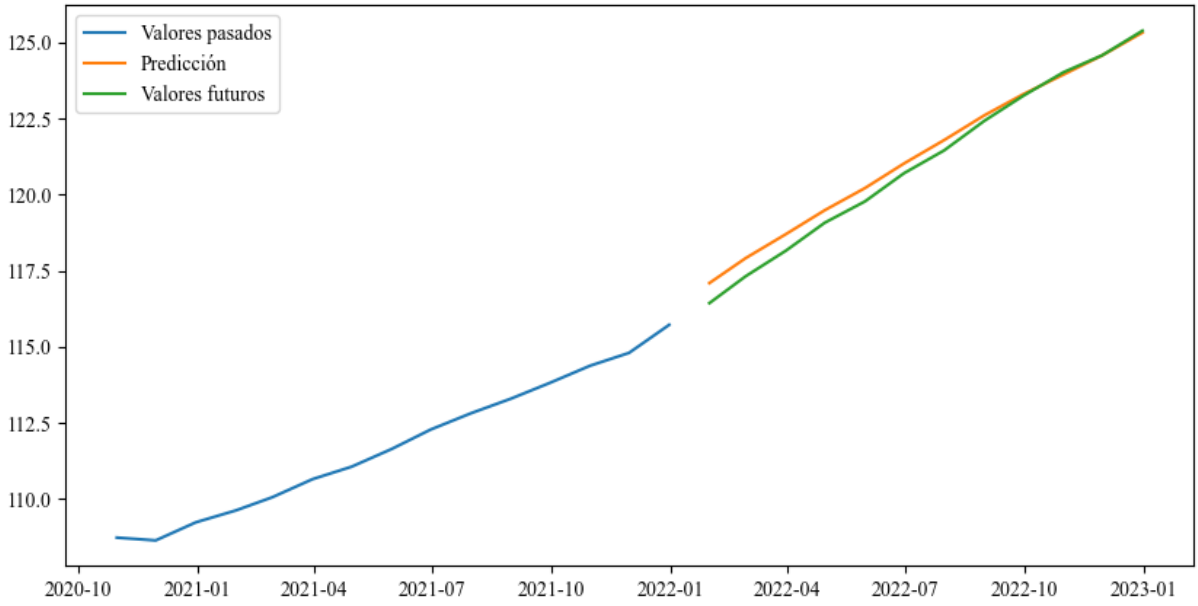
Figura 6: Calibración del modelo *LSTM* para pronosticar el INPC Subyacente



Fuente: elaboración propia con datos del SIE Banxico.

La Figura 7 corresponde a la predicción del INPC Subyacente en el periodo 2022, la cual muestra una convergencia entre el pronóstico y la serie real en los últimos meses de 2022.

Figura 7: Predicción del INPC Subyacente en el periodo 2022 mediante el modelo *LSTM*

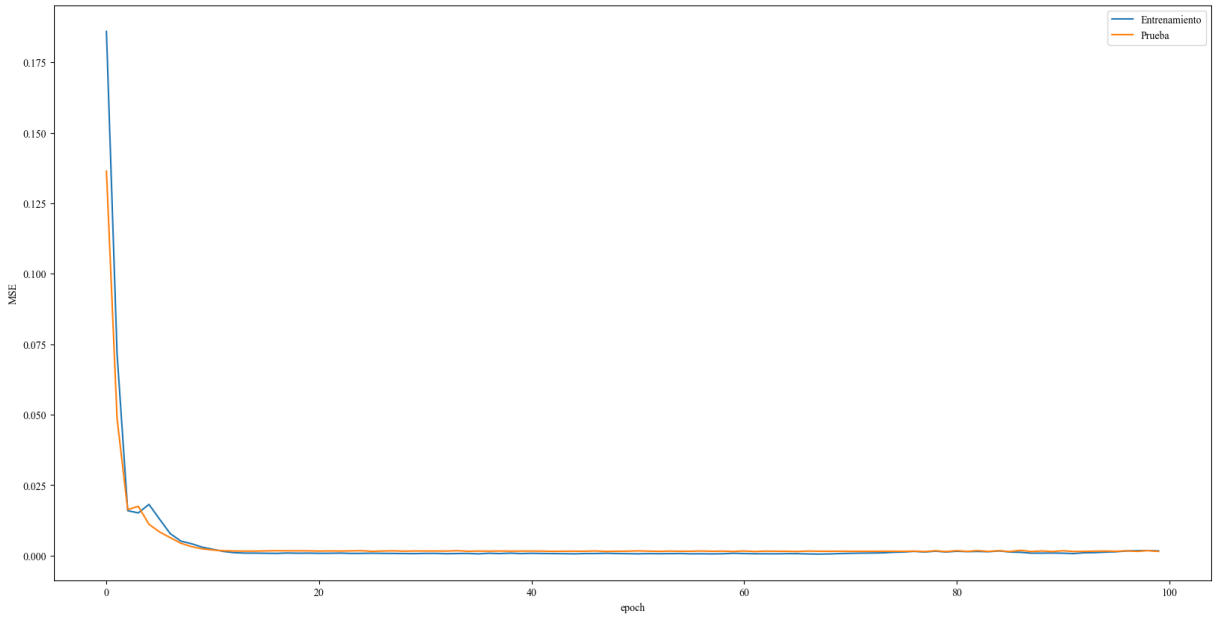


Fuente: elaboración propia con datos del SIE Banxico.

8.1.1.3. INPC No Subyacente

La figura 8 muestra la calibración del modelo *LSTM* para pronosticar el INPC no subyacente. Cabe destacar que el modelo no presentó sobreajuste y aprendió los patrones de la serie después de la época 7. Una época representa un ciclo de entrenamiento con los datos de entrenamiento y de prueba.

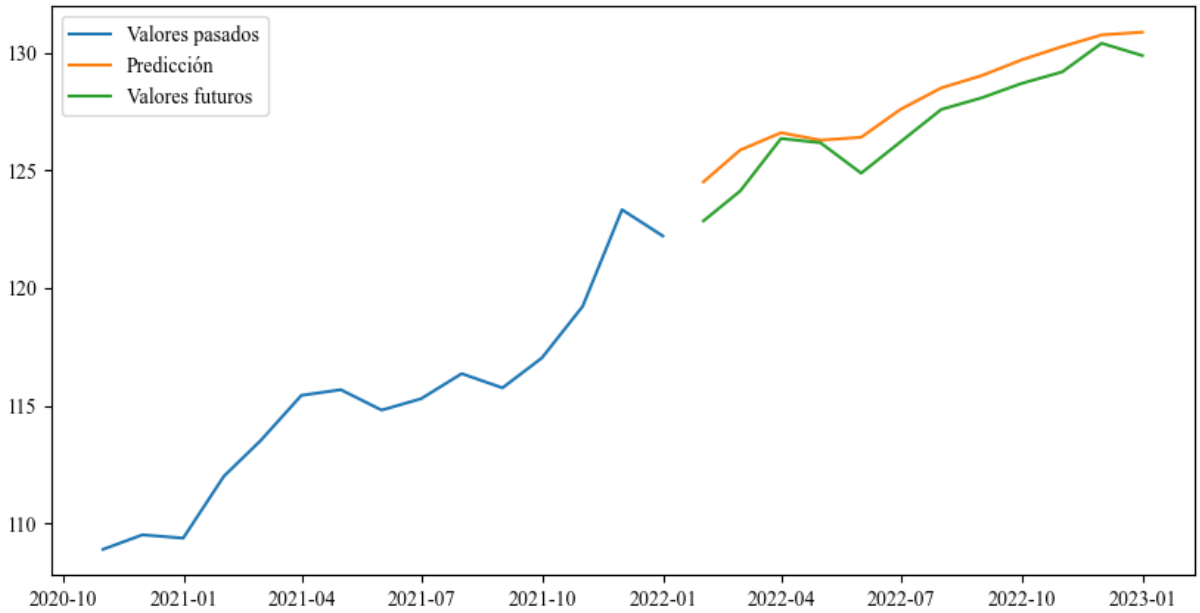
Figura 8: Calibración del modelo *LSTM* para pronosticar el INPC No Subyacente



Fuente: elaboración propia con datos del SIE Banxico.

La Figura 9 representa la predicción del INPC en el periodo 2022. Podemos notar que el modelo representa de forma suavizada los saltos que la serie real presentó durante 2022.

Figura 9: Predicción del INPC No Subyacente en el periodo 2022 mediante el modelo *LSTM*



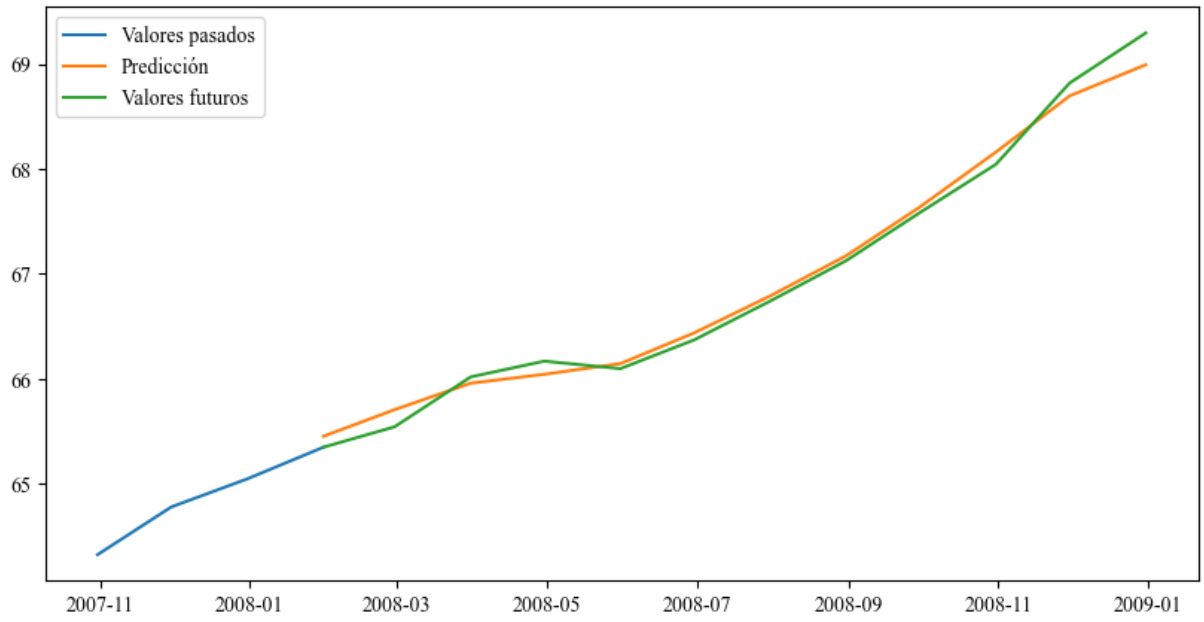
Fuente: elaboración propia con datos del SIE Banxico.

8.1.2. Predicciones 2008

8.1.2.1. INPC General

La Figura 10 representa la predicción del INPC general durante el periodo 2008. Es de importancia destacar que el pronóstico mostró un alto grado de convergencia con la serie real excepto para el salto de diciembre de 2008.

Figura 10: Predicción del INPC General en el periodo 2008 mediante el modelo *LSTM*

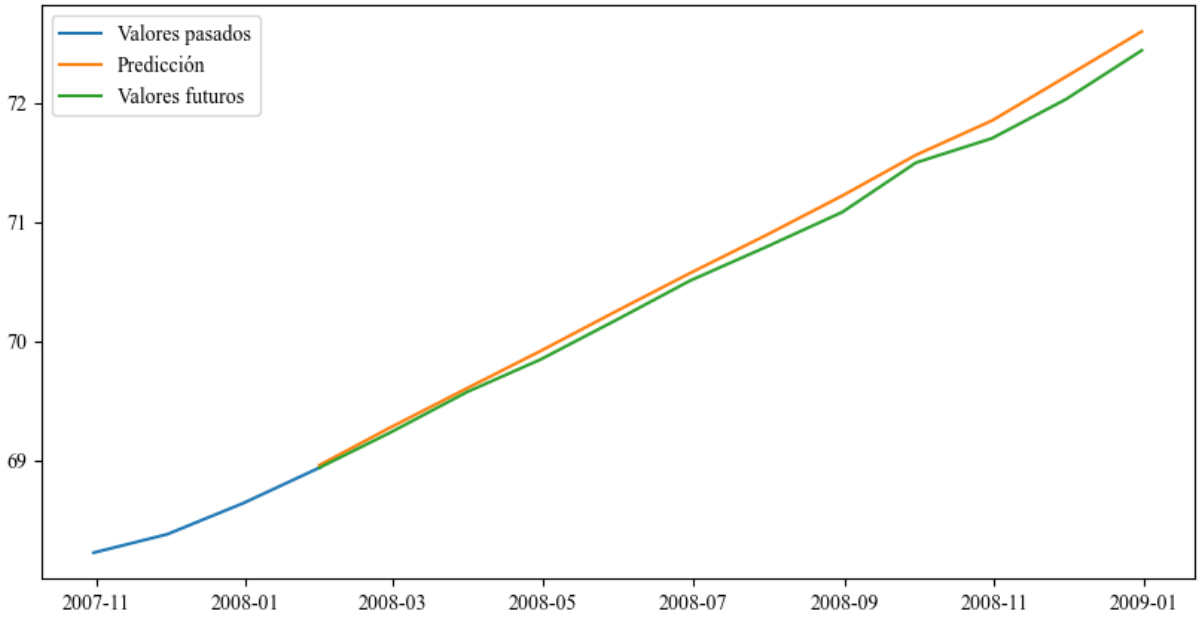


Fuente: elaboración propia con datos del SIE Banxico.

8.1.2.2. INPC Subyacente

La figura 11 ilustra la predicción del INPC subyacente para el periodo 2008. Podemos ver que el pronóstico del modelo representó de forma fiel el comportamiento de la serie durante dicho periodo.

Figura 11: Predicción del INPC Subyacente en el periodo 2008 mediante el modelo *LSTM*



Fuente: elaboración propia con datos del SIE Banxico.

8.1.2.3. INPC No Subyacente

Podemos ver la predicción del INPC no subyacente para el periodo 2008 en la Figura 12. El modelo presentó un pronóstico con expectativa alcista.

Figura 12: Predicción del INPC No Subyacente en el periodo 2008 mediante el modelo *LSTM*

Fuente: elaboración propia con datos del SIE Banxico.

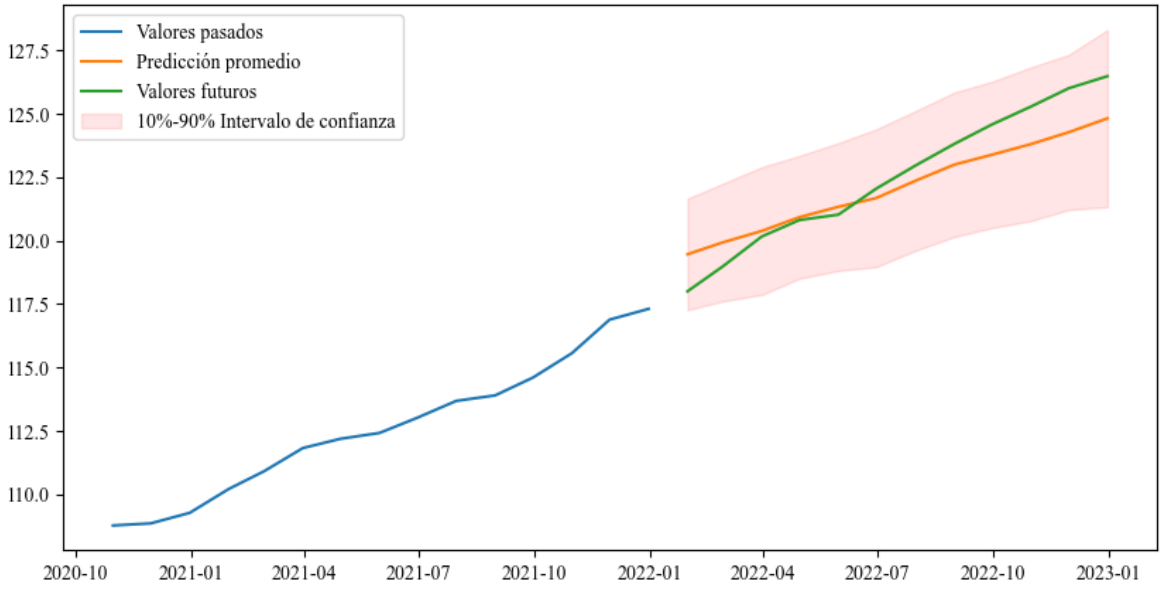
8.2. Pronósticos del modelo Auto*ML*

8.2.1. Predicciones 2022

8.2.1.1. INPC General

La Figura 13 muestra la predicción del INPC general en el periodo 2022. Los modelos de Auto*ML* producen pronósticos probabilísticos, por lo que el contorno de color anaranjado claro representa los cuantiles de la distribución del pronóstico (10%-90%). Así mismo, la línea anaranjada representa la media de las predicciones (cuantil 0.5).

Figura 13: Predicción del INPC General en el periodo 2022 mediante el modelo Ensamble-Auto*ML*

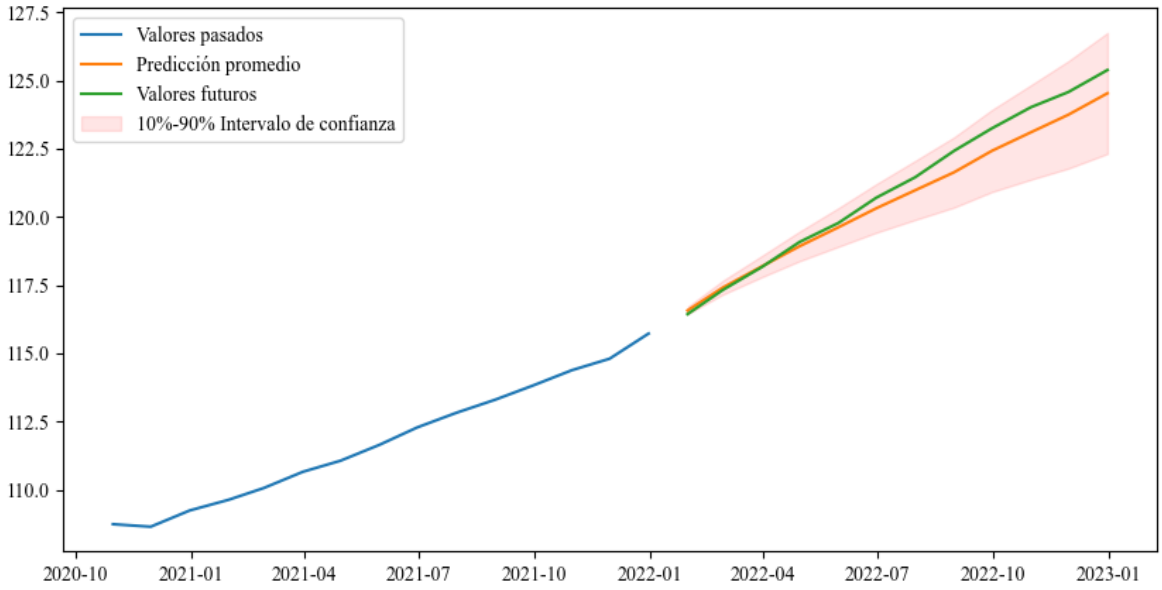


Fuente: elaboración propia con datos del SIE Banxico.

8.2.1.2. INPC Subyacente

La Figura 14 muestra la predicción del INPC subyacente en el periodo 2022. Los modelos de Auto*ML* producen pronósticos probabilísticos, por lo que el contorno de color anaranjado claro representa los cuantiles de la distribución del pronóstico (10%-90%). Así mismo, la línea anaranjada representa la media de las predicciones (cuantil 0.5).

Figura 14: Predicción del INPC Subyacente en el periodo 2022 mediante el modelo ETS-Auto*ML*



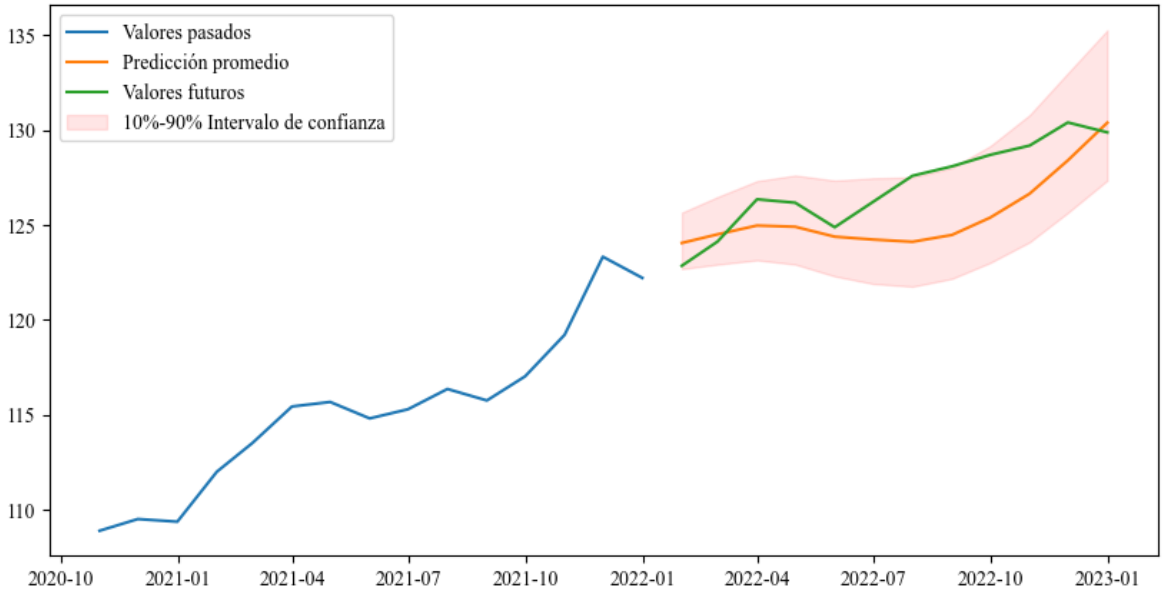
Fuente: elaboración propia con datos del SIE Banxico.

8.2.1.3. INPC No Subyacente

La Figura 15 muestra la predicción del INPC no subyacente en el periodo 2022. A pesar de que los valores de la serie real están dentro del intervalo de confianza de la predicción, podemos notar que el modelo pronostica un comportamiento convexo en la serie para la mitad del año, lo cual parece contrario al comportamiento real de la serie.

Los modelos de Auto*ML* producen pronósticos probabilísticos, por lo que el contorno de color anaranjado claro representa los cuantiles de la distribución del pronóstico (10%-90%). Así mismo, la línea anaranjada representa la media de las predicciones (cuantil 0.5).

Figura 15: Predicción del INPC No Subyacente en el periodo 2022 mediante el modelo Temporal Fusion Transformer-Auto*ML*



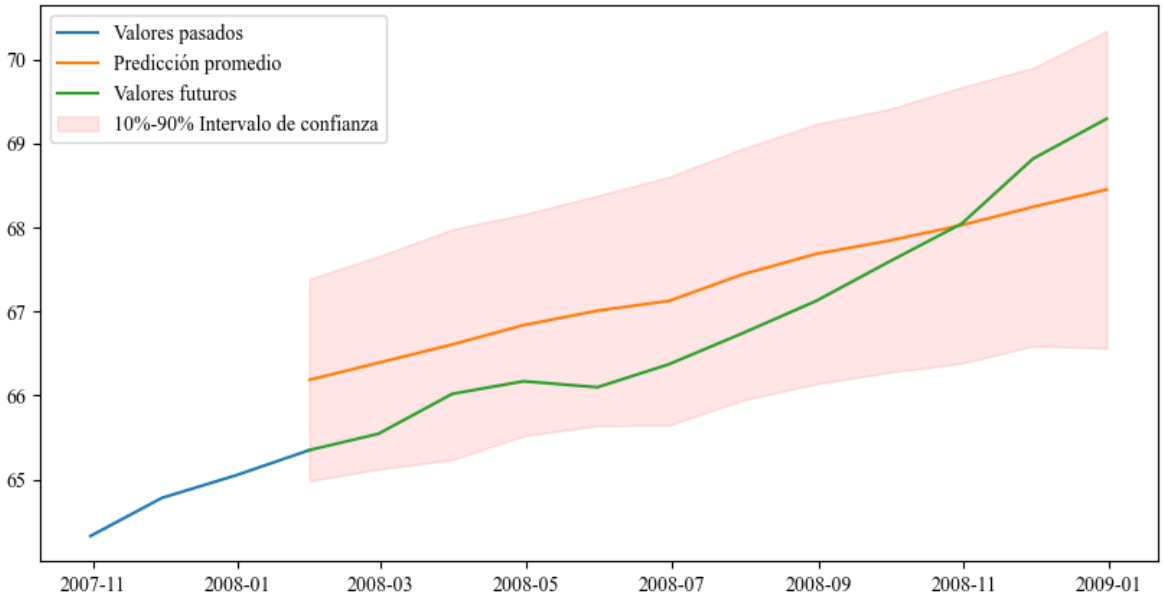
Fuente: elaboración propia con datos del SIE Banxico.

8.2.2. Predicciones 2008

8.2.2.1. INPC General

La Figura 16 muestra la predicción del INPC general en el periodo 2008. Los modelos de Auto*ML* producen pronósticos probabilísticos, por lo que el contorno de color anaranjado claro representa los cuantiles de la distribución del pronóstico (10%-90%). Así mismo, la línea anaranjada representa la media de las predicciones (cuantil 0.5).

Figura 16: Predicción del INPC General en el periodo 2008 mediante el modelo Ensamble-Auto*ML*

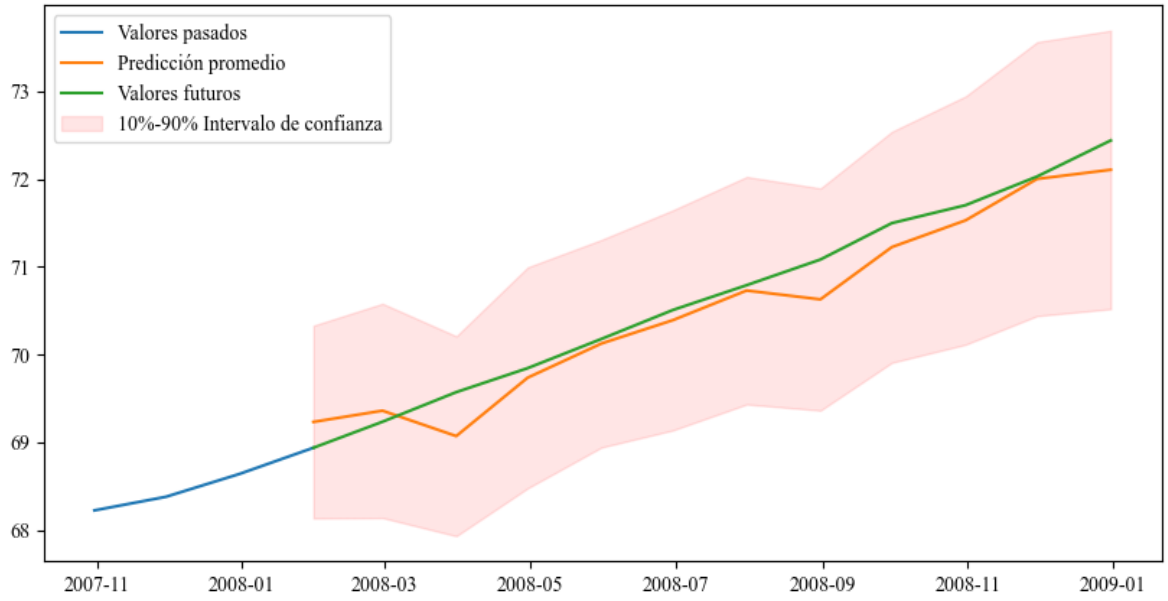


Fuente: elaboración propia con datos del SIE Banxico.

8.2.2.2. INPC Subyacente

La Figura 17 muestra la predicción del INPC subyacente en el periodo 2008. Los modelos de Auto*ML* producen pronósticos probabilísticos, por lo que el contorno de color anaranjado claro representa los cuantiles de la distribución del pronóstico (10%-90%). Así mismo, la línea anaranjada representa la media de las predicciones (cuantil 0.5).

Figura 17: Predicción del INPC Subyacente en el periodo 2008 mediante el modelo ETS-Auto*ML*

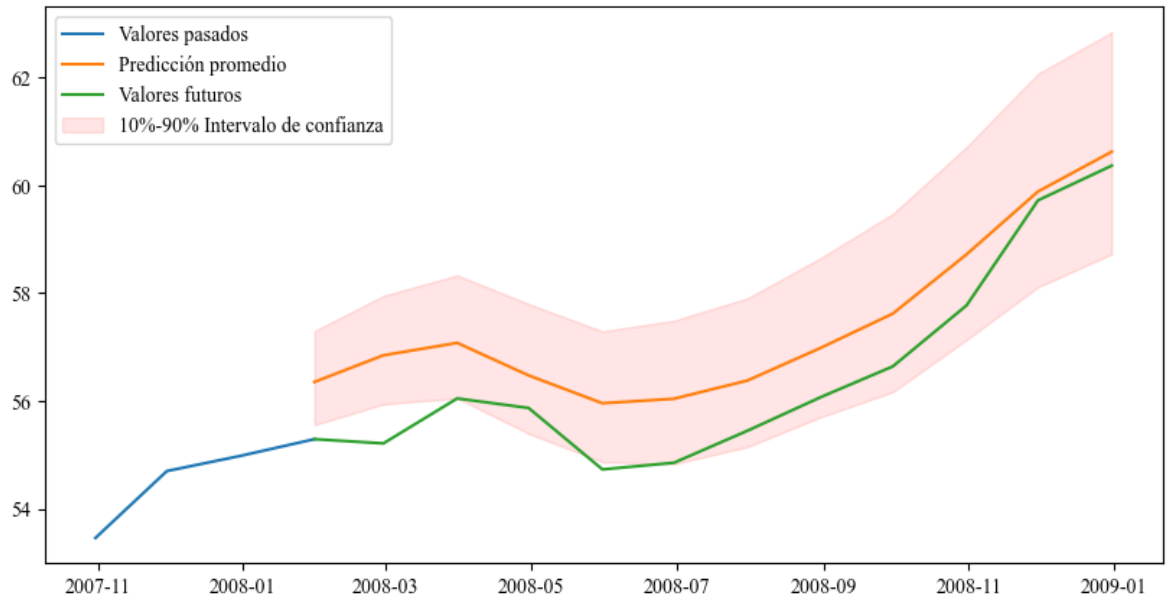


Fuente: elaboración propia con datos del SIE Banxico.

8.2.2.3. INPC No Subyacente

La Figura 18 muestra la predicción del INPC no subyacente en el periodo 2008. Los modelos de Auto*ML* producen pronósticos probabilísticos, por lo que el contorno de color anaranjado claro representa los cuantiles de la distribución del pronóstico (10%-90%). Así mismo, la línea anaranjada representa la media de las predicciones (cuantil 0.5).

Figura 18: Predicción del INPC No Subyacente en el periodo 2008 mediante el modelo Temporal Fusion Transformer-Auto*ML*



Fuente: elaboración propia con datos del SIE Banxico.

8.3. Pronósticos del modelo AR

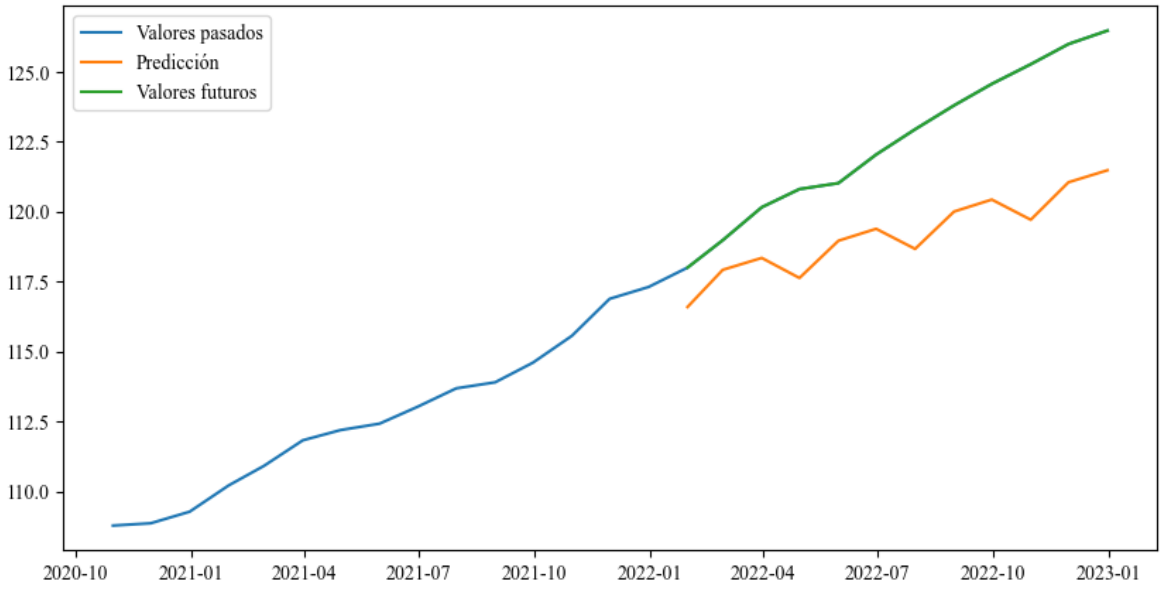
8.3.1. Predicciones 2022

8.3.1.1. INPC General

La Figura 19 muestra la predicción del INPC general en el periodo 2022, en la cual podemos observar que la brecha entre la predicción y la serie real se fue acrecentando hasta diciembre de 2022. Esto se debe a que la serie presentó un mayor grado de volatilidad debido al conflicto entre Rusia y Ucrania.

*Cabe destacar que el modelo AR es de orden (3,2), determinado por el criterio AIC.

Figura 19: Predicción del INPC General en el periodo 2022 mediante el modelo AR



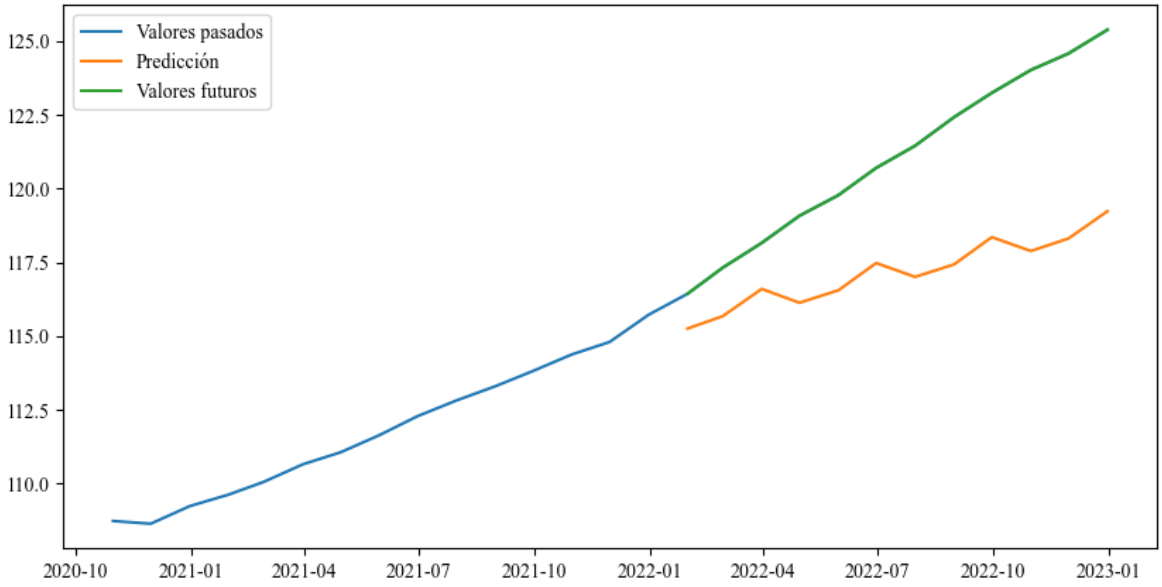
Fuente: elaboración propia con datos del SIE Banxico.

8.3.1.2. INPC Subyacente

La Figura 20 muestra la predicción del INPC subyacente en el periodo 2022, en la cual podemos observar que la brecha entre la predicción y la serie real se fue acrecentando hasta diciembre de 2022. Esto se debe a que el componente subyacente resultó particularmente afectado por el incremento en precios de alimentos debido al conflicto entre Rusia y Ucrania.

*Cabe destacar que el modelo AR es de orden (3,2), determinado por el criterio AIC y BIC.

Figura 20: Predicción del INPC Subyacente en el periodo 2022 mediante el modelo AR

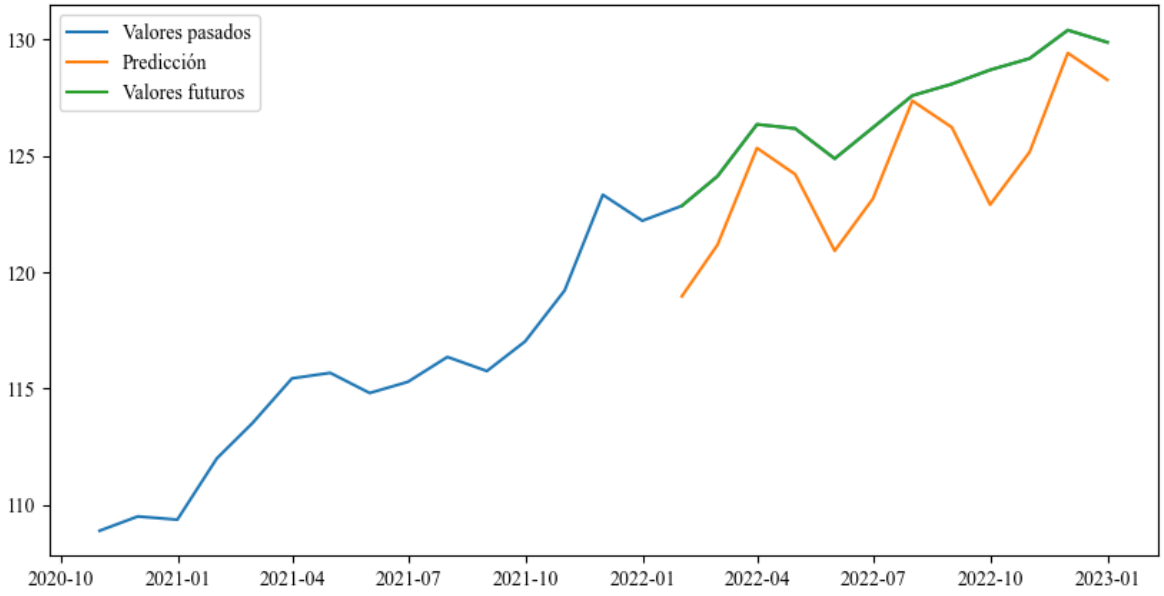


Fuente: elaboración propia con datos del SIE Banxico.

8.3.1.3. INPC No Subyacente

La Figura 21 muestra la predicción del INPC no subyacente en el periodo 2022. Cabe destacar que el modelo AR es de orden (4,2), determinado por el criterio AIC.

Figura 21: Predicción del INPC No Subyacente en el periodo 2022 mediante el modelo AR



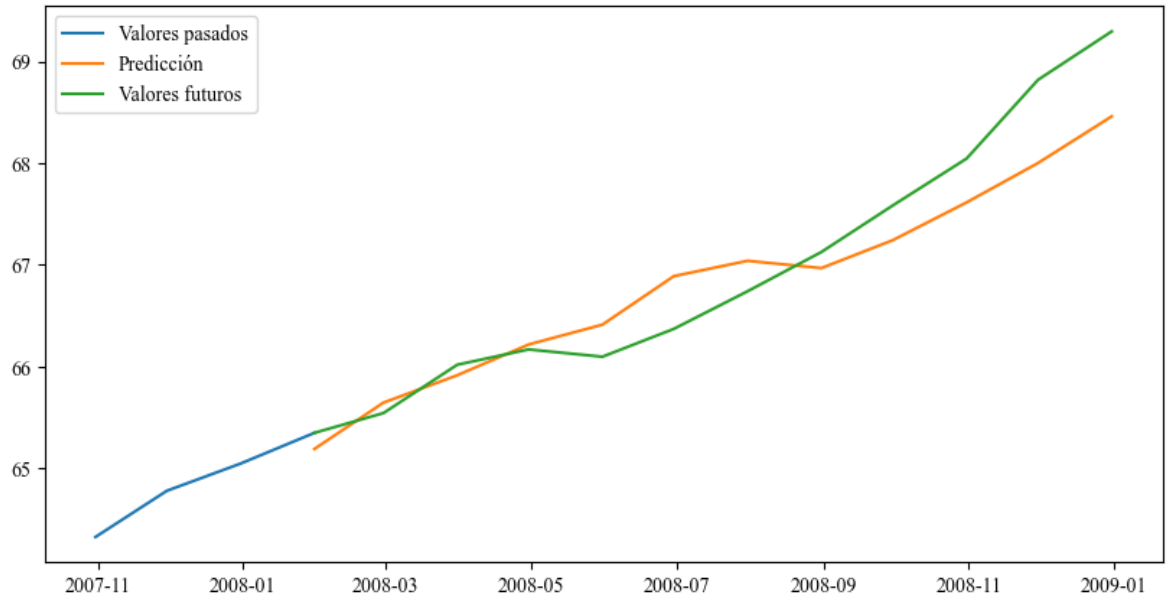
Fuente: elaboración propia con datos del SIE Banxico.

8.3.2. Predicciones 2008

8.3.2.1. INPC General

La Figura 22 muestra la predicción del INPC general en el periodo 2008. Cabe destacar que el modelo AR es de orden 3, determinado por el criterio AIC.

Figura 22: Predicción del INPC General en el periodo 2008 mediante el modelo AR

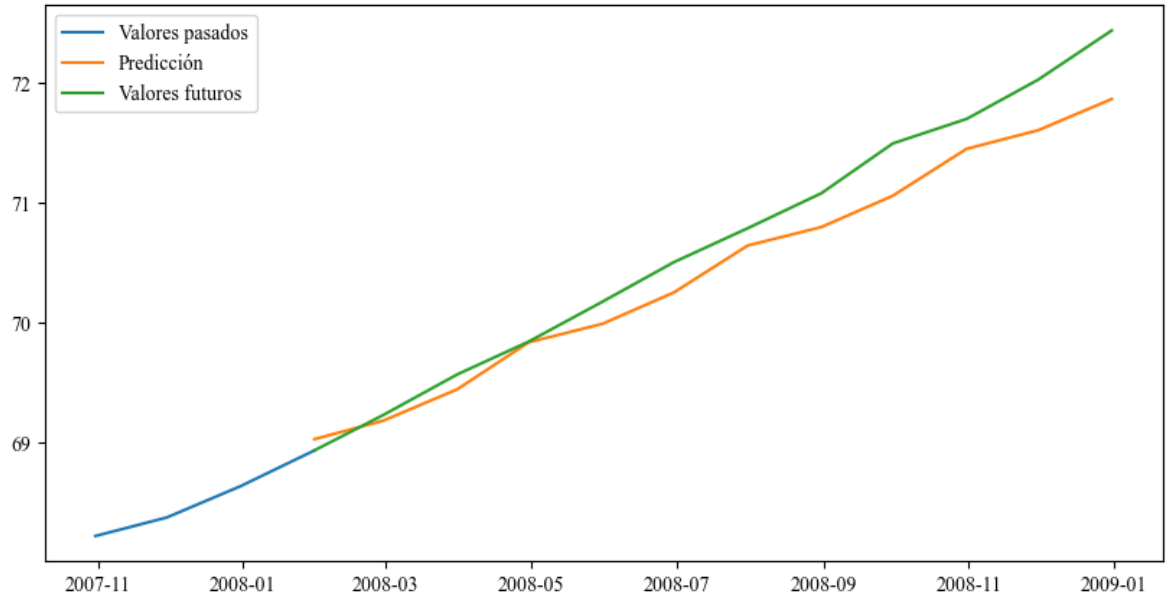


Fuente: elaboración propia con datos del SIE Banxico.

8.3.2.2. INPC Subyacente

La Figura 23 muestra la predicción del INPC subyacente en el periodo 2008. Cabe destacar que el modelo AR es de orden (3,2), determinado por el criterio AIC y BIC.

Figura 23: Predicción del INPC Subyacente en el periodo 2008 mediante el modelo AR

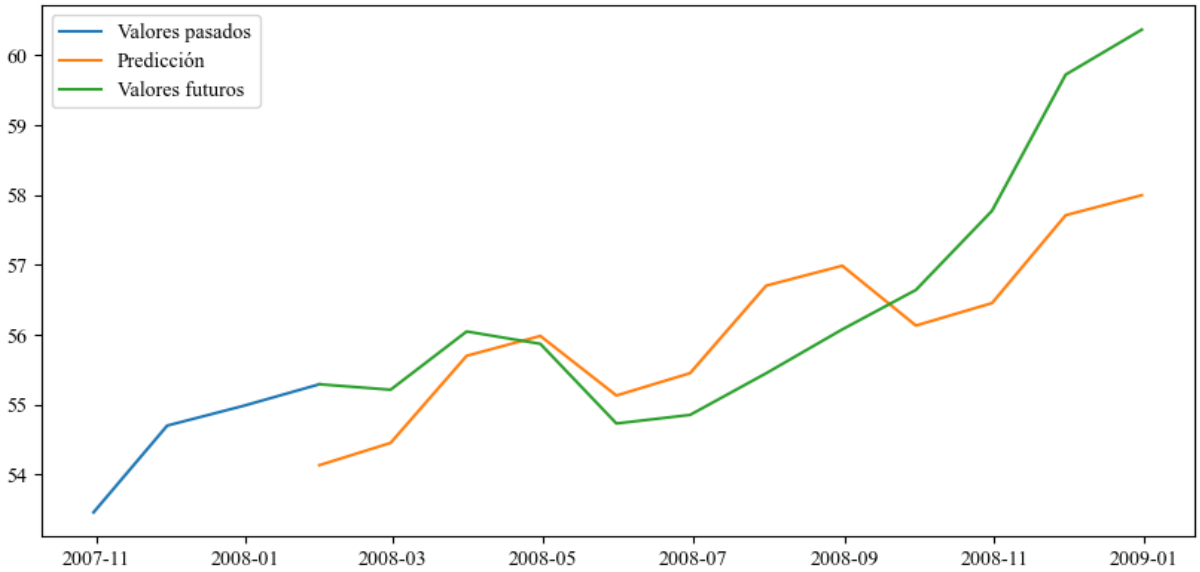


Fuente: elaboración propia con datos del SIE Banxico.

8.3.2.3. INPC No Subyacente

La Figura 24 muestra la predicción del INPC no subyacente en el periodo 2008. Cabe destacar que el modelo AR es de orden (4,2), determinado por el criterio AIC.

Figura 24: Predicción del INPC No Subyacente en el periodo 2008 mediante el modelo AR



Fuente: elaboración propia con datos del SIE Banxico.