

Benemérita Universidad

Autónoma de Puebla

Facultad de Economía

Secretaría de Investigación y Estudios de Posgrado

"APLICACIÓN DE MODELOS DE SERIES DE TIEMPO MULTIVARIABLES Y DE INTELIGENCIA ARTIFICIAL PARA ESTIMACIÓN DE PIB POTENCIAL EN MÉXICO"

T E S I S

Que para obtener el grado de
Maestro en ECONOMÍA.

Presenta(n):
Javier Tolentino Ramírez

Director(a) de tesis
Adrián Jiménez Gómez

Comité Tutorial:
* Alberto Castañón Herrera
*Roberto Olmos Pimentel

Puebla, Pue. 30 de agosto de 2024

Dedicatoria

Dedico esta tesis a mis padres y hermano, quienes sentaron las bases de mi motivación por el aprendizaje, responsabilidad y constancia, al Dr. Manuel Isidro Martin Ortiz a quien admiro y es un gran guía, a las contadas personas, que por confidencialidad no puedo poner sus nombres, pero que han sido un fuerte soporte emocional escuchando mis desventuras y aventuras, a mis colegas, amigues y todas las personas que han estado presente en este viaje, gracias a todos, porque me considero un conjunto de todas sus experiencias y las llevo presente.

Agradecimientos

Agradezco al CONAHCYT por el apoyo brindado, a mi asesor, Dr. Adrián Jiménez Gómez que mostró sensibilidad académica y humana para la continuidad y finalización de este proyecto (con todo el aprendizaje en el área económica que ello implica), a mi co-asesor el Dr. Roberto Olmos Pimentel por su apoyo y motivación a incorporarme al mundo de la inteligencia artificial, al Dr. Alberto Castañón Herrera por motivarme a realizar el estudio de posgrado, a la Dra. Silvia Beatriz Guillermo Peón por su dedicación a la enseñanza y su maravillosa labor como docente de estadística, a la coordinación y demás profesores de la facultad que hicieron significativo el aprendizaje en estos dos valiosos y complejos años.

Contenido

Dedicatoria 1

Agradecimientos 2

Índice de tablas, gráficas, diagramas e ilustraciones 5

Resumen 8

Introducción 9

Definición del objeto de estudio y su delimitación espacio-temporal 9

Justificación e implicaciones de la investigación 9

Identificación básica del objeto a investigar 10

Planteamiento del problema. 10

Objetivos generales y particulares 12

Hipótesis 13

Capítulo 1. Las lecciones del análisis económico 14

1.1. Los inicios del sistema económico 14

1.1.1. El sistema formalizado 16

1.1.2. El problema del PIB potencial 17

1.2. Los modelos autorregresivos en el análisis del PIB 18

1.3. La inteligencia artificial 20

Capítulo 2. Entre la econometría y la inteligencia artificial 25

2.1. Selección del modelo económico 25

2.2. Los modelos econométricos 28

2.2.1. El modelo VAR y VEC 30

2.3. Los modelos de inteligencia artificial 31

2.3.1. Redes Neuronales recurrentes y la red LSTM 36

2.3.2. Los optimizadores de la red neuronal 39

Capítulo 3. Los modelos y sus parámetros 42
3.1. Recopilación de la información y limpieza de datos 42
3.2. Aplicación del modelo VAR y VEC 44
3.3. Aplicación del modelo de Inteligencia Artificial 45
Capítulo 4. Resultados 47
4.1. Análisis 67
Conclusiones 70
Bibliografía 72
Apéndice 79

Índice de tablas, gráficas, diagramas e ilustraciones

Tablas

Tabla 2-1: Comparación de elementos de una neurona biológica y una neurona artificial 33

Tabla 2-2**:** Tipos de funciones de activación de una neurona artificial 35

Tabla 2-3: Detalle de los pasos del algoritmo ADAM 40

Tabla 3-1: Resultados pruebas Dickey-Fuller aumentadas para las variables en niveles 47

Tabla 3-2: Resultados pruebas Dickey-Fuller aumentadas para las variables en primeras diferencias 47

Tabla 3-3: Prueba de cointegración de Gregory-Hansen en modelo 3, cambio de nivel con tendencia 48

Tabla 3-4: Sistema VAR, máximo orden de retardos 9 49

Tabla 3-5: Prueba de Causalidad de Granger para VAR / Prueba de Wald de exogeneidad 50

Tabla 3-6: Pruebas de normalidad residual VAR con ortogonalización Cholesky (Lutkepohl) .. 50

Tabla 3-7: Pruebas de normalidad residual VAR con ortogonalización de correlación residual (Doornik-Hansen) 51

Tabla 3-8: Pruebas de heterocedasticidad residual VAR (niveles y cuadrados) 51

Tabla 3-9: Pruebas de heterocedasticidad residual VAR por componentes individuales 51

Tabla 3-10: Pruebas de LM de correlación serial residual VAR por rezago 52

Tabla 3-11: Pruebas de LM de correlación serial residual VAR por rezago acumulado 52

Tabla 3-12: Prueba de Cointegración de Johansen 53

Tabla 3-13: Prueba de rango de cointegración sin restricciones (valor propio máximo) 53

Tabla 3-14: Pruebas de significancia de los parámetros de corrección y vector de cointegración 54

Tabla 3-15: Pruebas de normalidad residual VEC con ortogonalización Cholesky (Lutkepohl). 56

Tabla 3-16: Pruebas de normalidad residual VEC con ortogonalización de correlación residual (Doornik-Hansen) 56

Tabla 3-17: Pruebas de heterocedasticidad residual VEC (niveles y cuadrados) 56

Tabla 3-18: Pruebas de heterocedasticidad residual VEC por componentes individuales: 57

Tabla 3-19: Pruebas de LM de correlación serial residual VAR por rezago 57

Tabla 3-20: Pruebas de LM de correlación serial residual VAR por rezago acumulado 58

Tabla 3-20: Valores de errores por número de rezago en red LSTM 64

Tabla 3-21: Porcentaje acumulado de la variación entre los resultados del PIB real y el PIB potencial para cada modelo 66

Gráficas

Gráfica 3-1: PIB estimado mediante VEC 59

Gráfica 3-2: PIB potencial estimado cuando la tasa de desempleo es 0 mediante VEC 59

Gráfica 3-3: PIB estimado cuando se incrementa el Acervo de Capital en un 10 % mediante VEC 60

Gráfica 3-4: PIB estimado cuando se incrementa el Acervo de Capital en un 10 % mediante VEC 60

Gráfica 3-5: PIB estimado mediante red LSTM en conjunto de entrenamiento 61

Gráfica 3-6: PIB estimado mediante red LSTM en conjunto de prueba 62

Gráfica 3-7: PIB potencial estimado cuando la tasa de desempleo es cero mediante red LSTM en conjunto de prueba 62

Gráfica 3-8: PIB potencial estimado cuando el Acervo de Capital se incrementa en 10% red LSTM en conjunto de prueba 63

Gráfica 3-9: PIB potencial estimado cuando el Acervo de Capital se incrementa en 20% red LSTM en conjunto de prueba 63

Gráfica 3-10: PIB potencial estimado cuando el desempleo se disminuye en 90% red LSTM en conjunto de prueba 65

Diagramas

Diagrama 2-1:Análisis estructural del PIB Potencial 26

Diagrama 2-2: Esquema de una red neuronal computacional 33

Diagrama 2-3: Representación de una Red Neuronal Recurrente 36

Diagrama 2-4: Funcionamiento de una neurona tipo LSTM 38

Ilustraciones

Ilustración 2-2: Representación de una neurona biológica 32

Resumen

El presente trabajo se centra en la estimación de una función producción, desde las perspectivas teórica y empírica que ofrecen los métodos econométricos de series multivariadas y los avances tecnológicos que ofrece la inteligencia artificial. Con ese fin se realiza un breve recorrido de las teorías económicas clásicas y de los problemas que estas enfrentan, dentro de los que se incluyen limitaciones para abordar los avances tecnológicos y los cambios económicos dinámicos.

La metodología de investigación propuesta incluye una revisión de modelos económicos históricos, análisis del PIB potencial mediante diversas técnicas, la aplicación de modelos VAR-VEC y de Inteligencia Artificial para la estimación económica. Lo último permite obtener las ventajas y desventajas de ambos modelos estadísticos que son considerados avanzados dentro del campo económico.

Introducción

Definición del objeto de estudio y su delimitación espacio-temporal

La proyección de PIB potencial puede realizarse mediante diversas técnicas estadísticas. Dentro de las herramientas disponibles para el pronóstico de variables económicas existen los modelos econométricos y los modelos de inteligencia artificial. Se plantea que los modelos de inteligencia artificial son más eficaces que los modelos econométricos. Para evaluar dicho planteamiento se propone proyectar el PIB potencial mediante ambas herramientas disponibles.

Debido a la importancia que tiene el análisis del PIB potencial en los análisis de crecimiento y desarrollo es de sumo interés realizar la proyección para el caso de México. A su vez, el periodo seleccionado es a partir de 1993, un año antes de la crisis de 1994, desde la cuál el país se ha encontrado con estabilidad relativa de precios, hasta el periodo de información disponible de las variables explicativas PIB en México, que para el presente trabajo es el primer trimestre del 2023.

Justificación e implicaciones de la investigación

Las series de tiempo multivariable representan un gran avance en la estimación de modelos dinámicos dentro de la economía, lo cual es de gran utilidad para incorporar a los análisis econométricos la naturaleza cambiante e interrelacionada de sus variables. Uno de los modelos más destacados en ese campo es el de Vectores Autorregresivos y de Corrección de Errores.

Por otra parte, las redes neuronales son capaces de trabajar información que no es linealmente separable (NLS) lo que permite que sus tareas de clasificación y regresión sean fácilmente adaptadas a problemas económicos (Herbrich et al., 2012), de la definición de dichos problemas nos centramos dos aplicaciones principales:

1. Clasificación de agentes económicos: Permite clasificar un conjunto de información NLS en diferentes categorías a diferencia de modelos tradicionales como análisis discriminatorio lineal o regresiones logísticas. Muchos de estos estudios se enfocan en la predicción de bancarrota de agentes económicos, principalmente en el sector bancario.
2. Predicción de series de tiempo: Probablemente la aplicación más reconocida del uso de Redes Neuronales en la economía se encuentra en la predicción de mercado de capitales. Los resultados de la aplicación de estos nuevos métodos computacionales han demostrado una mayor eficiencia en el desarrollo de pronósticos bursátiles.

El poder aplicar dichos modelos al área económica permitirá desarrollar nuevas técnicas que estén adaptadas a la innovación teórica y práctica mediante el uso del software disponible. Los resultados obtenidos pueden ser utilizados por instituciones públicas como privadas para mejorar la toma de decisiones en relación con las políticas fiscales y monetarias del país. Otro aspecto importante es que ampliaría el alcance de herramientas especializadas a un sector más grande, para futuras investigaciones en diversos campos donde sea relevante la estimación.

Identificación básica del objeto a investigar

En la proyección de variables, los modelos de inteligencia artificial han demostrado altos niveles de eficiencia para diversos campos y áreas de estudio. Los primeros hallazgos sobre la eficiencia de estos modelos se han encontrado desde su aplicación a problemáticas de las ciencias naturales.

Las aplicaciones más comunes de la inteligencia artificial se encuentran en los trabajos de clasificación (para detección automática de imágenes), más recientemente se popularizó para lenguaje natural (Chat GPT3) y en el ámbito económico se popularizó para series de tiempo, en específico para el área financiera.

Para evaluar la validez de los modelos económicos se debe trabajar con información ya conocida, es decir, a diferencia de otras ciencias no es posible aislar los agentes económicos para analizar su comportamiento. Por tanto, es de suma importancia tener modelos estadísticos eficientes para realizar correctas interpretaciones de los modelos económicos.

La eficiencia que han mostrado los modelos de inteligencia artificial para analizar información conocida y cuyos resultados se desarrollan en diversos campos del conocimiento, es un fuerte indicador de la pertinencia que tiene la incorporación de los modelos de inteligencia artificial a la evaluación empírica de los modelos económicos.

Planteamiento del problema.

Los modelos para estimar el PIB potencial se dividen en aquellos que interpretan a dicha métrica como un valor de equilibrio de la producción, en caso de encontrarse por encima de dicho valor la tendencia es disminuya en los próximos periodos hasta regresar a su equilibrio y viceversa, en otras palabras, este primer enfoque se refiere a una tendencia del nivel de crecimiento de un país. La segunda perspectiva hace alusión a una interpretación más directa, es decir, refiere al PIB

potencial como una métrica de un nivel máximo de producción que es alcanzable sólo en el supuesto de plena utilización de sus factores de producción.

Además de dicha división sintética también existe una división en los tipos de modelos a utilizar para estimar el PIB potencial, los cuales se refieren específicamente a los modelos de series multivariables, que son relativamente recientes en el área económica, y los modelos de inteligencia artificial que han mostrado un auge en los últimos años gracias a la disponibilidad de la información y programas de cómputo para diversas áreas del conocimiento.

Respecto al último punto podemos decir que los modelos tradicionales para simular pronósticos en el área económica presentan limitaciones y reducen su confiabilidad debido a varios factores. Evans (1999) realiza un trabajo preciso al momento de clasificar las problemáticas de la aplicación de modelos económicos a la vida real, o gran parte de ellas, mediante tres limitantes:

a) Supuestos incorrectos: Dado que los modelos se basan en premisas que no siempre resultan ciertas, por ejemplo, en política monetaria se puede asumir que las instituciones financieras deciden de forma eficiente para promover el desarrollo económico mientras que otros autores pueden demostrar que estas instituciones actúan en su propio beneficio.
b) Exceso de simplificación: Los modelos macroeconómicos son ampliamente usados debido a que permiten simplificar situaciones escogiendo sólo un número de factores que parecen tener el mayor impacto en la proyección a analizar omitiendo todos los demás.
c) Modelos matemáticos demasiado complejos: La enorme cantidad de variables y ecuaciones que buscan modelar algún fenómeno económico representan un problema a la hora de resolver dichos sistemas matemáticamente, lo que lleva generalmente a linealizarlos de forma que sea más sencillo resolverlos, sin embargo, la economía no exhibe patrones necesariamente lineales.

El manejo del primer problema se observa mediante la evaluación empírica de las diversas teorías del pensamiento económico, es decir, las herramientas estadísticas han sido suficientes para identificar las imprecisiones de los modelos teóricos, mediante su evaluación empírica. En otras palabras, la creación de modelos matemáticos más eficientes aporta marginalmente poco a la primera limitación que plantea Evans.

El exceso de simplificación tiene dos vertientes, la primera es el debate acerca del uso excesivo de modelos estadísticos en la investigación económica, dándole más importancia a los resultados numéricos por encima del análisis teórico del modelo económico. La segunda vertiente es sobre la

utilidad de identificar los aspectos más relevantes de causalidad, evitando tener un número excesivo de variables que marginalmente, *ceteris paribus*, aportan poco a la explicación del fenómeno. En relación con ambas vertientes existen argumentos para utilizar datos menos agregados, microeconómicos, para las estimaciones empíricas de modelos económicos (Felipe y McCombie, 2007, 226).

Antes de abordar la tercera limitante cabe destacar que dentro de los modelos computacionales existen las llamadas redes neuronales. Éstas pueden concebirse como un modelo matemático creado con base en un gran número de elementos organizados en niveles y consiste en un número determinado de neuronas artificiales clasificadas en tres niveles principales: entradas, capas ocultas y salidas (Panchal y Panchal, 2014, 456-457). Dichos modelos buscan una aproximación al modelo biológico del cerebro humano, considerando que cada neurona computacional busca imitar el comportamiento general de una neurona biológica y por tanto no replica de forma exacta cada una de las funciones de esta última.

Dada la complejidad de las variables que componen un modelo económico, éstas suelen linealizarse como una forma de afrontar dicha complejidad, facilitando con ello el análisis matemático en el que se apoya el modelo, sin embargo, el comportamiento económico no es lineal y muchas características se pierden dicha linealización.

En ese sentido la última limitante puede superarse gracias a las habilidades de las redes neuronales de trabajar con información no linealmente separable, lo que permite que sus tareas de clasificación y regresión sean fácilmente adaptadas a problemas económicos.

En ese orden de ideas, es posible inferir que las estimaciones de variables económicas obtenidas mediante modelos de inteligencia artificial tienden a ser más eficientes que aquellas relacionadas a modelos estadísticos tradicionales, aportando soluciones a las problemáticas planteadas para el pronóstico de variables económicas. Lo último sin dejar de lado las posibles limitaciones de dichos modelos y las ventajas que tienen los modelos de series multivariadas en su interpretación económica.

Objetivos generales y particulares

a) GENERALES
   - Identificar el nivel del PIB que debemos de alcanzar a través de la utilización plena de todos los factores de la producción disponibles en México.

b) PARTICULARES

- Proponer un modelo para el cálculo del PIB potencial, congruente con la teoría económica y rigurosa en su parte estadística.
- Aplicar el modelo propuesto mediante Vectores Autorregresivos y Vector de Corrección de errores para la estimación del PIB potencial en México.
- Identificar los alcances y las limitaciones de la aplicación del modelo propuesto, mediante Vectores Autorregresivos y Vector de Corrección de errores, considerando las características macroeconómicas de México.
- Realizar un modelo de inteligencia artificial para proyectar el PIB potencial en México.
- Identificar los alcances y las limitaciones del modelo de inteligencia artificial considerando las características macroeconómicas de México.

Hipótesis

Los modelos de inteligencia artificial permiten estimar el PIB potencial de manera más precisa que el modelo de Solow aumentado, estimado mediante estimadores de Vectores Autorregresivos y Vector de Corrección de errores.

El primer capítulo de la tesis realiza un recorrido por la literatura existente sobre el crecimiento económico, el uso de la matemática para los análisis empíricos, así como la incorporación de la Inteligencia Artificial al análisis económico.

En el segundo capítulo se selecciona el modelo lineal más adecuado para la estimación del PIB y del PIB potencial, así como el modelo de inteligencia artificial que se ajuste mejor a la predicción de series temporales.

En el tercer capítulo se detalla la metodología utilizada para la aplicación empírica de los modelos propuestos, incluyendo la búsqueda, limpieza y utilización de los datos.

En el cuarto capítulo se muestran los resultados de la aplicación de los modelos propuestos y posteriormente se realiza un análisis de dichos resultados incluyendo las conclusiones correspondientes para la presente y futuras investigaciones.

Capítulo 1. Las lecciones del análisis económico

1.1. Los inicios del sistema económico

Los modelos económicos buscan representar la realidad del contexto social y cultural en el que los seres humanos intercambian bienes y servicios para subsistir, definiendo a "las economías" como la estructura de cada país o el conjunto de países que comparten una base económica similar. Por cuestiones prácticas nos enfocaremos en las teorías que más impacto tienen en el actual sistema económico, entendiéndose al capitalismo en el sentido de Jahan y Mahmud (2015) como el sistema base de casi todas las economías del mundo.

Si bien, la definición de capitalismo entra constantemente en debate debido a los diferentes métodos para interpretarlo, la variedad de modelos económicos establece un conjunto de supuestos que pueden o no ajustarse a cada definición de capitalismo, o en su defecto, a la representación de una economía en específico. La diferencia entre estos se escapa al alcance del presente trabajo.

Para representar dicha realidad los primeros teóricos de la economía moderna, también conocidos como "los clásicos"[1], sentaron las bases de la producción y la demanda, identificando los aspectos fundamentales que intervienen en el proceso de intercambio de valor y de acumulación de riqueza.

Smith (1784) identificó el importante papel de la especialización en la producción de las empresas, que se traduciría en rendimientos marginales crecientes del capital, a pesar de reconocer que los bienes son escasos. Un avance destacable fue reconocer en forma indirecta el papel de la tecnología, en la acumulación de riqueza de una nación, como un factor endógeno del crecimiento.

Por otra parte, David Ricardo y Robert Malthus establecieron lo que después sería conocido como el "pesimismo". Puesto que ellos vivieron en una época en donde la población había crecido a un ritmo acelerado, consideraron que los incrementos en la producción se verían limitados por los recursos disponibles. En palabras de Malthus, "La población, cuando no es controlada, crece en una proporción geométrica. La subsistencia incrementa sólo en proporción aritmética" (Malthus, 1798, 4). Lo que no consideraron es que la tecnología de los años venideros sería capaz de superar esta limitante con creces.

[1] Adam Smith, David Ricardo, Thomas Robert Malthus, Karl Heinrich Marx.

Jean Baptiste Say fue reconocido por su aportación a la naturaleza cíclica de la oferta y la demanda, restando importancia al dinero en el proceso de intercambio de valor y haciendo énfasis en la producción. De su análisis es posible identificar una serie de conclusiones que se centran en la calidad del producto, la extensión del mercado, la abundancia de los bienes entre otras variables que tengan una incidencia positiva en la llamada capacidad productiva, llegando a aseverar que los buenos gobiernos se enfocan en mejorar la productividad y que los malos gobiernos en incrementar el consumo (Say, 1880).

Cabe destacar que Karl Marx había previsto con anterioridad en su famoso texto, El Capital, lo que llamó "las contradicciones del capitalismo". La más importante, en relación con la crisis de 1929, es el doble carácter del trabajo, el cual refiere a la forma natural y a la forma de valor que se les da a las mercancías. No hay que olvidar que el desplome financiero se debió a una sobreproducción, generando lo que podría denominarse un exceso de "valor natural".

La crisis financiera de 1929 mostró que el libre mercado no era suficiente para la estabilidad económica. Por ello, Keynes (1936) realizaría un análisis que le otorgó nuevamente importancia al gobierno como un ente regulador, haciendo particular énfasis en las deficiencias que el libre mercado trajo consigo.

En contraposición a la idea de Keynes de establecer un modelo de equilibrio estático Harrod (1939) se cuestionaba acerca de la naturaleza dinámica de la economía. El modelo que desarrolló le permitió identificar tres tasas de crecimiento: la tasa observada, la garantizada y la natural. La relación entre las dos primeras establecía el equilibrio en el corto plazo y las dos últimas en el largo plazo. Los mismos resultados fueron conseguidos por una vía diferente por parte de Domar (1947).

Uno de los principales problemas del modelo, que después fue conocido como el modelo Harrod-Domar, era su alto grado de inestabilidad. Dicho problema intentó ser resuelto por Robert Merton Solow en su artículo "A Contribution to the Theory of Economic Growth" publicado en 1956. Trevor Swan, al igual que Domar, encontró de forma independiente los mismos resultados.

Cada modelo refleja parte de la realidad social e histórica en la que fue desarrollado además de hacer énfasis en el carácter exógeno o endógeno de las variables macroeconómicas. Las críticas y mejoras que le han sido realizadas a cada modelo económico se relacionan directa o indirectamente con la falta de precisión al evaluarlos empíricamente.

1.1.1. El sistema formalizado

A partir de la década de 1950 existió un giro en el enfoque metodológico de la teoría económica, lo anterior debido al debate sobre su carácter como una disciplina formal. Lo cual llevó paulatinamente a una transición de la explicación literaria a las demostraciones matemáticas como base del pensamiento económico (Wulwick, 1992).

A pesar de que ya existían modelos con un fuerte sesgo cuantitativo, como el caso del modelo de Solow-Swan, dicha transición afectó indistintamente a todos los modelos que se escribirían en años posteriores, y en consecuencia, al tema de las nuevas aportaciones, enfocándose principalmente en el aparato matemático.

Otra de las consecuencias destacables de dicha transición fue la incorporación de los modelos estadísticos para realizar proyecciones de las variables macroeconómicas. Es importante destacar que las técnicas econométricas han sido criticadas, por ejemplo, McCloskey (1985) nos menciona la tendencia de los economistas en confiar demasiado en los resultados estadísticos por sí mismos, dejando de lado el análisis general del modelo.

En ese sentido Gordon Tullok atribuye a Ronald Coase la frase "Si torturas lo suficiente los datos, ellos confesarán", haciendo referencia a las "modificaciones" en las especificaciones del modelo econométrico para producir resultados que sean más significativos en términos estadísticos (Klein y Tullok, 2001, 83-87).

Replicar los resultados también representa un problema en los estudios que utilizan herramientas estadísticas, un tópico en el que Tullok difiere con Daniel Bruce Klein, en alusión al término de estudios "rigurosos". Los resultados publicados en estudios con valoraciones estadísticas deberían ser replicables, Levy y Feigenbaum (1990) realizaron una evaluación de los coeficientes obtenidos por otros estudios y encontraron discrepancias que pueden deberse a que: los datos se mezclan después de que los coeficientes correctos son publicados o los datos son correctos y los coeficientes reportados no lo son.

De lo anterior puede deducirse que la transición de la economía literaria a la economía como ciencia se ha enfrentado a una serie de dificultades que se han sorteado de diferentes maneras. La economía es en ese sentido, una ciencia joven e incompleta, por ello es importante considerar estas desventajas en los estudios económicos empíricos para reducir el margen de error o el riesgo de producir resultados espurios.

1.1.2. El problema del PIB potencial

El PIB potencial generalmente se refiere a la capacidad de producción máxima de una economía sin provocar presiones inflacionarias. Es un concepto teórico que estima el nivel de actividad económica al que se alcanza el pleno empleo de los recursos disponibles en la economía, incluyendo mano de obra y capital, sin generar aumentos en la inflación. Dicho de otra forma, el PIB potencial es útil para evaluar la capacidad y fortaleza económica de un país, así como para identificar brechas del producto, es decir, la diferencia entre el PIB observado y el PIB potencial. Usualmente dicha diferencia es positiva o negativa, lo cual sugiere si una economía está por debajo o sobre su nivel de producción óptimo, lo cual puede resultar contradictorio si consideramos que el PIB potencial se refiere a un máximo.

Podemos citar algunos trabajos que se han realizado respecto a la estimación del PIB Potencial, por ejemplo, Corona et al. (2022) realiza un estudio del PIB Potencial utilizando cuatro técnicas: métodos heurísticos, el filtro de Hodrick Prescott, el Modelo de Factores Dinámicos no estacionario y la descomposición Permanente-Transitorio (PT) de Gonzalo y Granger (1995).

Acevedo Rueda et al. (2012) utilizan en el desarrollo de su estimación un modelo de Vectores Autorregresivos (VAR) de cuatro rezagos. En dicho modelo incluyeron algunas variables dicotómicas para capturar efectos específicos como altos niveles de inflación, crisis petroleras y financieras.

Chen y Papanyan (2016) utilizan el enfoque de la función de producción, dicho enfoque es preferido por organizaciones como la Oficina de Presupuesto del Congreso (CBO), la Comisión de la Unión Europea y la Organización para la Cooperación y el Desarrollo Económicos (OCDE). Este método da como resultado lo que se denomina PIB "potencial de plena capacidad", lo que sugiere un límite superior para el desempeño económico y delinea las "brechas de producción" resultantes como más amplias y profundas que las derivadas de las estimaciones del potencial tendencial. A pesar de que los autores no especifican el modelo utilizado, brindan pistas sobre el camino que debe seguirse para la estimación del PIB potencial.

Los trabajos de estimación realizado al respecto se pueden clasificar en dos grupos principales, aquellos que entienden al PIB potencial como una tendencia del crecimiento económico, manteniendo ciertas condiciones de estabilidad, y aquellos que lo interpretan, desde una perspectiva más directa, como aquella capacidad productiva máxima, la cual en estricta

definición no es estadísticamente alcanzable. Cabe destacar que la mayor proporción de investigaciones realizadas utilizan el primer enfoque.

Debido a la falta de acuerdo sobre la definición del PIB potencial, no existe una metodología precisa para calcularlo, lo que conduce a los estudios relacionados a enfocarse en el ámbito exploratorio. En ese sentido la Oficina de Presupuesto del Congreso menciona que existe un espectro de opiniones entre los economistas acerca de la utilidad de las medidas del PIB potencial para la política monetaria y fiscal y para las proyecciones económicas, haciendo alusión a que algunos economistas no creen que la idea de que la producción potencial sea útil bajo argumentos como:

> "[…]
>
> El PIB potencial es tan inestable y varía tanto que es imposible estimarlo con precisión, especialmente en los últimos años, por lo que no es una guía útil para la formulación de políticas o la previsión.
>
> Por lo general, las políticas de gestión de la demanda hacen más daño que bien debido a los retrasos, las incertidumbres y las presiones políticas. Por lo tanto, el tamaño de la brecha entre la producción real y la potencial debería ser irrelevante para los responsables de la formulación de políticas." (The Congress of the United States, 2004, p.2)

Dichas aseveraciones se sustentan principalmente en los métodos actuales para su cálculo, ya que es difícil medir algo que no tiene un consenso general. Sin embargo, las actuales herramientas estadísticas, computacionales y de disponibilidad de información están dando un giro a la actual tendencia de los estudios sobre las variables macroeconómicas. En el presente trabajo se busca incorporar dichos elementos para proponer un modelo para el cálculo del PIB potencial, congruente con la teoría económica y rigurosa en su parte estadística.

1.2. Los modelos autorregresivos en el análisis del PIB

Los modelos de vectores autorregresivos (VAR) consideran la naturaleza dinámica del comportamiento económico, por lo cual se han utilizado una gran cantidad de variaciones de estos modelos para estimar y proyectar variables macroeconómicas. A pesar de la particularidad que buscamos, el nivel máximo del PIB, existe más literatura que refiere al uso de modelos VAR en términos generales del nivel del PIB que de su nivel potencial, motivo por el cual se hace un breve recuento de los trabajos realizados al respecto sin importar si refieren a su nivel observado o potencial.

Roush et al. (2017) utilizaron un modelo VAR (4), donde 4 es el número de rezagos, construido para el PIB real de EE. UU. El modelo mostró signos de funcionamiento adecuado con datos históricos, pero ofrece información predictiva limitada, pronosticando un crecimiento económico lineal hasta el primer trimestre de 2018, en ese sentido recomiendan un marco de modelización más avanzado para predecir mejor los movimientos futuros del PIB.

Pilström y Pohl (2009) realizaron un trabajo de tesis que tuvo como objetivo establecer un modelo de previsión del crecimiento del PIB en los Estados bálticos mediante un modelo VAR reducido, usando como variables los valores pasados del PIB, la inflación y el desempleo son variables explicativas del modelo. Las conclusiones de los autores mencionan que el modelo elegido proporciona previsiones fiables del PIB de los Estados bálticos y que se puede aplicar a otros países. Un par de aspectos son destacables de dicha investigación, la primera es que la volatilidad debido a los cambios institucionales, aunado a los cambios económicos de la región, disminuyen significativamente la capacidad predictiva del modelo VAR reducido y la segunda es que los horizontes mayores a 12 cuatrimestres son altamente inciertos y poco confiables.

Anaya Narváez et al. (2016) utilizaron un modelo VAR para analizar el crecimiento económico colombiano tomando como base teórica un modelo extendido de Solow, que incorpora el capital humano como factor determinante. Los autores incluyeron un Modelo de Corrección de Errores, después de mostrar evidencia de cointegración entre las variables de series temporales. Sus resultados soportaron la teoría del capital humano en el crecimiento económico para Colombia.

Existen una gran cantidad de trabajos que utilizan los modelos de análisis de series de tiempo, en específico los modelos VAR y VEC, así como algunas derivaciones de éstos. Los resultados en general son positivos y en su mayoría coinciden con una limitante principal, relacionada a la estabilidad de las variables en el sentido económico, es decir, los modelos que utilizan el concepto de cointegración requieren de una combinación lineal de parámetros que relacionan las variables de interés con un comportamiento estocástico en común, de modo que los choques o cambios estructurales presentan más de una relación de cointegración, generando la necesidad de establecer diferentes modelos de corto y largo plazo para cada periodo con cambio estructural.

1.3. La inteligencia artificial

La inteligencia artificial (IA) desempeña un papel creciente en el análisis económico, abarcando áreas como previsión, modelado, toma de decisiones, comercio e inversión. En dicha área existen diversas técnicas de aprendizaje automático, por ejemplo, el procesamiento del lenguaje natural, que se utiliza para analizar datos de texto.

Estudios de diversas instituciones internacionales sugieren que la IA puede duplicar tasas de crecimiento económico global para 2035. Por ejemplo, la International Telecommunication Union. (2018) describe el impacto económico de la IA y su potencial para impulsar el crecimiento global. Analiza el estado actual de la tecnología de IA y sus aplicaciones en diversos sectores, destacando su capacidad para mejorar la productividad y generar PIB adicional. Sugiere que la IA podría contribuir con alrededor del 16% a la producción mundial para 2030, lo que equivale a 13 billones de dólares, pero también reconoce los riesgos potenciales y las externalidades negativas asociadas con la implementación de la IA.

Por su parte Ernst et. al. (2018) aborda la aceleración en el cambio de habilidades necesarias en la fuerza laboral debido a la automatización y la inteligencia artificial. Según el modelo de habilidades de la fuerza laboral del Instituto McKinsey Global, se predice un cambio en las horas trabajadas y en las habilidades requeridas en Estados Unidos y Europa Occidental para 2030. Mencionan que estos cambios afectarán tanto a las habilidades físicas y manuales como a las habilidades cognitivas, emocionales y tecnológicas, resaltando la importancia de entender las implicaciones económicas y sociales de la inteligencia artificial para el futuro del trabajo.

Dicho crecimiento se espera gracias a los avances que la IA promete en la productividad, creando fuerzas laborales virtuales y fomentando la innovación. Se espera que sectores como marketing, ventas, cadena de suministro, logística y fabricación se beneficien de la IA. Su impacto se extiende a diversas áreas como asistentes inteligentes, vehículos autónomos y autoselección de contenido para redes sociales.

Es importante mencionar que el auge de la IA se impulsa por avances en potencia informática, conectividad y disponibilidad de grandes volúmenes de datos, transformando sociedades y economías. Por ejemplo, China destaca en avances notables en iniciativas de IA tanto a nivel inversión como producción y aplicación.

En ese tenor existe un creciente interés por la IA y sus implicaciones, Damasevicius (2023) proporciona una visión general del uso de la inteligencia artificial (IA) en diversos campos, como

la economía, las finanzas, la minería, la fabricación y la innovación. Analiza los impulsores y los efectos del despliegue de la IA en el contexto de la innovación y se destacan los retos y las oportunidades que se derivan de su uso.

Sin embargo, también existen preocupaciones respecto al uso de la IA, en especial lo que respecta a la sobreestimación de sus beneficios y la subestimación de sus costos reales, los cuales a su vez son difíciles de rastrear o cuantificar. La última premisa no es sorpresa si consideramos que la cadena de producción que lleva al servicio final de la IA es demasiado amplia e incluye un fuerte peso en la parte de servicios, principalmente de propiedad intelectual.

En ese orden de ideas podemos citar a Szczepański (2019), quien explora preocupaciones como el efecto disruptivo de la IA en la economía y la sociedad, la ampliación de la desigualdad y el desplazamiento de puestos de trabajo. O el trabajo de Agrawal et al. (2019), quienes mencionan que existe una falta de medidas adecuadas para evaluar el rendimiento y la mejora del desarrollo de la IA, en el sentido de la falta de medidas de crecimiento de la productividad, cuando este crecimiento es impulsado principalmente por el capital intangible. Otros problemas que se plantean son la forma de medir los beneficios de la IA y su impacto en la desigualdad.

Tomando en consideración que la relación costo beneficio de la IA es debatible, la creación del conocimiento y la evaluación de nuevas técnicas no lo es, por ello es importante tomar el presente estudio con el cuidado que se merece, en el sentido de no idealizar a la IA como un solucionador de problemas universales, ya que los mismos han demostrado deficiencias que sólo al análisis tradicional puede afrontar.

A continuación, se mencionarán estudios y métodos utilizados para el pronóstico de variables económicas con uso de la IA y su núcleo, que en el presente estudio son las redes neuronales.

i. Kaastra y Boyd (1995) analizan el diseño de una red neuronal para pronosticar datos de series temporales financieras y económicas. Destacan el mayor uso de redes neuronales en finanzas para el reconocimiento de patrones, la clasificación y la predicción de series temporales. Dichos autores utilizan un caso específico de redes neuronales de retropropagación (BP).
ii. Gallardo Del Angel (2020) analiza la previsión de series temporales financieras mediante redes neuronales artificiales (RNA). Describe el análisis utilizando el algoritmo tradicional de retropropagación y el algoritmo de retropropagación resiliente. Su estudio se centra en la predicción de los movimientos de los precios de las acciones y evalúa la importancia

relativa de las diferentes variables características utilizadas para la predicción. Los datos utilizados para el análisis incluyen seis índices bursátiles principales de Asia, Europa y América del Norte.

iii. Huang et. al. (2007) se enfocan en tres áreas principales: pronóstico de tasas de cambio, pronóstico de precios de índices bursátiles y pronóstico de crecimiento económico. Analizan diversos modelos de redes neuronales, como los perceptrones multicapa (MLP), las redes neuronales recurrentes (RNN) y las redes neuronales probabilísticas (PNN). Destacan la importancia de integrar las redes neuronales con otras tecnologías, como algoritmos genéticos, análisis de wavelet, inferencia difusa y reconocimiento de patrones, para mejorar el rendimiento de las redes neuronales en la predicción financiera y económica

iv. Faria y Leibovici (2023) exploran el uso de Modelos de Lenguaje Grande (LLM) para generar pronósticos de inflación condicionales. Los autores comparan los pronósticos generados por PaLM LLM de Google AI con la Encuesta de Pronosticadores Profesionales (SPF), encontrando que los pronósticos de LLM tienen errores cuadráticos medios más bajos en general. Los autores sugieren que los LLM pueden proporcionar pronósticos baratos y precisos, aplicables a otras variables macroeconómicas. Los pronósticos de PaLM consideran factores como las tasas de interés, el desempleo y los precios del petróleo.

v. Goulet (2022) analiza un nuevo método basado en el aprendizaje automático, la Red Neuronal Hemisférica (HNN), para pronosticar y comprender la inflación. Compara el desempeño del modelo HNN con los modelos econométricos tradicionales en la predicción de la inflación en los Estados Unidos. El informe destaca la capacidad de HNN para capturar el reciente aumento de la inflación y su uso de indicadores alternativos como el índice de vacantes y las horas trabajadas para medir la holgura económica. El autor concluye que el aprendizaje profundo se puede utilizar para construir pronósticos interpretables económicamente, pero debe combinarse con el juicio de expertos.

La complejidad en el diseño de modelos de predicción de redes neuronales destaca la necesidad de prueba y error debido a la gran cantidad de parámetros involucrados. Por ello se sugiere mejorar el uso de redes neuronales en economía mediante el desarrollo de la teoría para manejar datos atípicos e incompletos. Si bien es cierto, las redes neuronales son una extensión de los modelos de

regresión con la ventaja de ser comparables con la organización del cerebro compuesto por neuronas interconectadas.

Para el caso específico de México, la literatura sobre la aplicación de modelos de IA al área económica es limitada y suelen usarse otros países o áreas para el análisis, por ejemplo, Sossa (2020) habla sobre el impacto de la IA desde el análisis de tecnologías como los *chatbots* para mejorar servicios sociales y *cobots*[2] que han optimizado procesos en la industria, como en la electrónica de consumo. En el ámbito educativo, narra soluciones innovadoras como la realidad virtual y sistemas de tutoría inteligente, que han facilitado el aprendizaje y la interacción de los estudiantes con la tecnología en México. Al respecto Albrieu et al. (2018) enfatizan el uso de la IA en la industria, proponiendo que dicha implementación permitiría incrementar la productividad, al tiempo que proponen un reto: la falta de adopción de tecnologías emergentes y el diseño de políticas públicas que fomenten el crecimiento y la equidad.

Otros autores como Valle-Cruz et al. (2023) identifican que la IA tiene el potencial de impactar significativamente a diversos sectores en México, particularmente en la presupuestación pública y la asignación de recursos, a través de una visión más profunda de los fenómenos sociales como variables complejas. Esto podría conducir a mejoras en el desarrollo económico, la reducción de la inflación y la mitigación de la desigualdad, ya que el análisis de datos exhaustivos permite tomar decisiones más informadas que prioricen las necesidades de la sociedad. Sin embargo, los autores reconocen que no sólo se requiere la información, también es necesaria una mejora en la gestión, lo cual implica volver más eficientes las decisiones dentro del aparato gubernamental.

La relevancia del uso de dicha herramienta computacional reside en que las predicciones macroeconómicas enfrentan desafíos debido a la falta de un modelo sólido y a la no estacionariedad de las series económicas. En el ámbito de la previsión financiera, se destacan obstáculos como el pequeño tamaño de las muestras, el comportamiento aleatorio de las series, la falta de estacionariedad y la presencia de no linealidad en los datos financieros.

Por poner un ejemplo, los modelos estadísticos convencionales en macroeconomía han resultado ineficaces para anticipar la persistencia de altas tasas de inflación recientes. Al grado que la rápida adopción de medidas de política monetaria restrictivas por parte de los bancos centrales

[2] Los *chatbots* son los asistentes virtuales que se encuentran en diversas páginas web de atención a clientes o mediante llamadas telefónicas que no tienen a una persona física realizando la actividad de contestar. Los *cobots* son asistentes en tareas de alto riesgo para los empleados, principalmente en la industria manufacturera.

ha suscitado preocupaciones sobre una posible recesión. El debate sobre la naturaleza transitoria o persistente de la inflación, junto con las dificultades para medir la diferencia entre la oferta y la demanda agregadas, plantea cuestionamientos cruciales en los ámbitos de la política monetaria y fiscal, los cuales se plantea en el presente estudio que pueden ser cuantificados de forma no lineal por las redes neuronales.

Gonzalez (2000) advierte que la literatura aún es inconclusa debido al número limitado de estudios confiables disponibles. Enumera las fortalezas y debilidades de las redes neuronales, como su capacidad para modelar relaciones no lineales, pero también la dificultad de interpretar los pesos estimados y la necesidad de muestras grandes. Comparó la red neuronal con un modelo de regresión lineal para pronosticar el crecimiento del PIB real de Canadá. Dicha red mostró una precisión de pronóstico superior tanto para los períodos dentro como fuera de la muestra. Sus pruebas estadísticas indican evidencia limitada de que la mejora en la precisión de los pronósticos sea estadísticamente significativa.

El último estudio citado destaca el desarrollo de los modelos de IA, los cuales han tenido un desarrollo importante en las últimas décadas, cuando se comparan los estudios realizados hace apenas 20 años encontramos bastantes deficiencias, tanto en el campo de la capacidad computacional como en las herramientas disponibles para sus cálculos, haciendo referencia específicamente a una menor capacidad de las computadoras y de los programas computacionales disponibles actualmente.

Capítulo 2. Entre la econometría y la inteligencia artificial

2.1. Selección del modelo económico

En México se utiliza el método de oferta agregada para el cálculo del PIB, es decir que se suma el consumo, la inversión, el gasto las exportaciones y las importaciones. Para obtener el total de cada rubro de valor agregado se realiza un análisis muestral por actividad productiva y se extrapola mediante técnicas econométricas, dicha muestra se toma de las 622 clases de actividad productiva que componen el Indicador Global de la Actividad Económica (IGAE).[3] Por su parte el PIB potencial se calcula, en México, como un porcentaje de crecimiento esperado dados los últimos 10 años, es decir, no como un límite superior sino como una estimación del PIB para cinco periodos posteriores[4].

Para el presente trabajo se utiliza el concepto del PIB potencial como un valor máximo o tope, bajo esa premisa es importante ofrecer un punto inicial con los estudios que han demostrado relación entre los factores de producción y el cálculo tradicional del PIB. En ese sentido es menester mencionar la Ley de Okun, la cual postula una relación inversa entre el tamaño de la brecha de producción (la diferencia porcentual entre el PIB y el PIB potencial) y el tamaño de la brecha de desempleo (la diferencia entre la tasa de desempleo y la tasa natural de desempleo).

Las ventajas de los modelos de oferta es la desagregación en factores productivos. Por ello, si consideramos el modelo económico más representativo para la producción nacional tenemos entonces una ecuación Cobb-Douglas, en donde Y representa la producción agregada, L es el insumo de trabajo, K es el insumo de capital, A es el nivel de tecnología y t es el tiempo en el que se realiza la medida de nivel de producción.

$$\boldsymbol{Y_t = A_t L^{\alpha}{}_t K^{\beta}{}_t} \qquad \text{ec. 1}$$

Tomando logaritmos naturales de la ecuación anterior obtenemos una relación de las elasticidades del factor de producción con respecto a la mano de obra y el capital α y β, respectivamente. Bajo el supuesto de rendimientos constantes la suma de las elasticidades debe ser igual a uno (α + β = 1).

$$\boldsymbol{\ln(Y_t) = \ln(A_t) + \alpha \ln(L_t) + \beta \ln(K_t)} \qquad \text{ec. 2}$$

[3] La metodología exacta puede consultarse en la página oficial del INEGI, en el apartado de Economía y Sectores Productivos.
[4] Artículo 11 C del Reglamento de la Ley Federal de Presupuesto y Responsabilidad Hacendaria

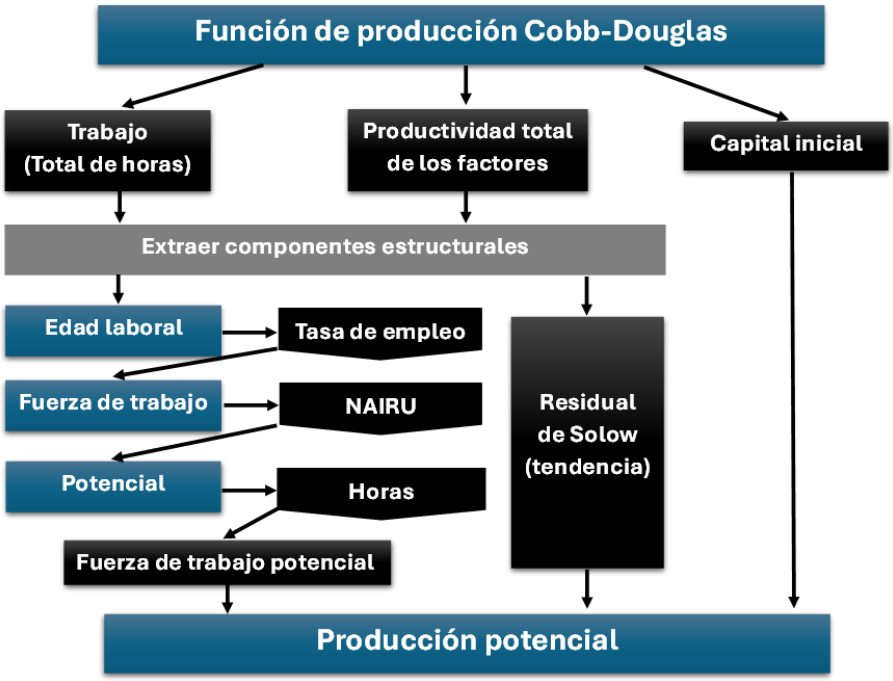
De esta última ecuación se deriva lo que se conoce como el Residual de Solow o la Productividad Total de los Factores (PTF), la cual se le atribuye directamente al nivel de tecnología y eficiencia agregado del sistema productivo.

$$\mathbf{PTF} = \ln(A_t) = \ln(Y_t) - \alpha \ln(L_t) - \beta \ln(K_t) \qquad \text{ec. 3}$$

De forma teórica, este residuo de Solow refleja la diferencia entre la producción real y un promedio ponderado de los insumos de trabajo y capital. Si el objetivo es obtener una estimación precisa de este residuo es esencial realizar una medición precisa de los insumos de trabajo y capital. Sin embargo, estudios como los de Felipe y McCombie (2007) mencionan una relación espuria de dicho coeficiente debido a la forma en que se mide el ingreso nacional. Lo último no implica que la productividad total de los factores sea incalculable, sólo hace referencia a los métodos tradicionales que se utilizan para calcularlo.

Por ejemplo, Chen y Papanyan (2016) utilizan una métrica global para el cálculo del PIB potencial, en el cual incorpora las variables que se utilizan tradicionalmente para el cálculo de la Ley de Okun y la PTF. Por razones de confidencialidad y propiedad intelectual no son extensos en su forma específica de calcularlo, sólo mencionan las variables que utilizan, a nivel general.

Diagrama 2-1:Análisis estructural del PIB Potencial



Gráfica de flujo modificada de Chen y Papanyan (2016)

Intuitivamente podría pensarse en el PIB potencial como una situación en la que se lleva el desempleo (mano de obra potencial) a una tasa 0 o, en el caso de que no exista capacidad ociosa, incrementar el capital productivo hasta que sea absorbido por la mano de obra existente. Sin embargo, dicho análisis se enfrenta a un problema fundamental encontrado en un conocido fenómeno económico, la Curva de Philips, la cual es una relación inversa entre la tasa de desempleo y la inflación, en otras palabras, disminuir la tasa de desempleo ocasiona un incremento en la tasa de inflación y viceversa.

La Curva de Philips es un fenómeno ampliamente analizado en la literatura económica, sin embargo, su relación con el crecimiento económico es usualmente tratada de forma indirecta, un ejemplo se encuentra en Clark y Laxton (1997) quienes mencionan que el desempleo excesivo como la alta inflación pueden obstaculizar el crecimiento económico. Un desempleo bajo junto con una inflación controlada puede significar una economía estable, promoviendo el gasto y la inversión de los consumidores. Asimismo, mencionan que las expectativas de inflación son un determinante crítico de la "razón de sacrificio", que se relaciona con el costo (en términos de desempleo) requerido para reducir la inflación. Esto implica que las expectativas de inflación bien gestionadas pueden dar lugar a entornos económicos más estables, fomentando una atmósfera propicia para el crecimiento. Esto indica el papel de la gestión de la inflación en el mantenimiento de un crecimiento económico estable, ya que el exceso de demanda y la inflación resultante pueden conducir a la inestabilidad económica.

Una crítica a la visión tradicional de la Curva de Philips la encontramos en Karanassou et al. (2006) quienes postulan que no existe una relación a largo plazo entre la inflación y el desempleo, es decir, que es una relación espuria. Este punto de vista se basa en la suposición de una curva de Phillips neo keynesiana vertical a largo plazo, dicho enfoque convencional supone la existencia de una tasa natural de desempleo (NRU, por sus siglas en inglés) y estima la tasa de desempleo no aceleradora de la inflación (NAIRU, por sus siglas en inglés) como la tasa de desempleo compatible con la estabilidad de la inflación, ignorando la posibilidad de que exista una disyuntiva entre la inflación y el desempleo. El argumento central es que la teoría ortodoxa mantiene supuestos inverosímiles, como una tasa de interés cero o una simetría de ponderación inter temporal, y no tiene en cuenta la persistencia sustancial de la inflación que puede surgir de la asimetría de dicha ponderación.

No es difícil deducir de lo anterior que el PIB potencial se ve limitado por la inflación a través de la interconexión dinámica existente con el desempleo, dicho de otro modo, se propone un modelo que pueda considerar la relación dinámica tanto a corto como a largo plazo de la persistencia de la inflación en un modelo de PIB potencial.

Es importante destacar que, los mecanismos de transmisión de las variables macroeconómicas varían según el contexto socio económico de cada país, en otras palabras, los supuestos de la economía heterodoxa consideran escenarios racionales, simétricos y en algunos variables valores constantes para su análisis, lo cual conlleva a resultados discrepantes cuando se incluyen más variables en los modelos. Tal es el caso de la búsqueda del PIB potencial, el cual debe entenderse desde el contexto de México.

2.2. Los modelos econométricos

Los modelos que tradicionalmente se han utilizado para la estimación de causalidad de variables son los modelos econométricos de regresión lineal, los cuales pueden ser simples o múltiples. De acuerdo a Bouza (2018):

> "El análisis de regresión es una técnica usada para modelar la relación entre variables. Se desea establecer como una o varias variables dependientes se comportan respecto a una o más variables independientes. Mediante esta técnica podemos obtener información sobre cómo una variable de interés Y, variable dependiente, reacciona cuando una de las independientes lo hace." (p. 6)

Un modelo de regresión lineal múltiple tiene la forma general

$$\boldsymbol{Y} = \mathbf{B}\boldsymbol{X} + \boldsymbol{\epsilon} \qquad \text{ec. 4}$$

Donde Y es un vector de dimensión N que representa a una variable continua o discreta, B es un vector de tamaño K con los parámetros representados por $\{\beta_0, \beta_1, \beta_2, \ldots, \beta_K\}$, los cuales deben ser estimados, X es una matriz de tamaño K x N de variables $\{X_0, X_1, X_2, \ldots, X_K\}$ que explican a Y, con $X_0 = \{1, 1, 1, \ldots, 1\}$, es decir, X_0 es un vector unitario de dimensión N, y ϵ es el vector de errores, también de tamaño N.

Para obtener los parámetros o estimadores de un modelo de regresión existen diversas técnicas, bajo un conjunto de premisas, conocidas como los supuestos del modelo de regresión lineal múltiple, la estimación por Mínimos Cuadrados Ordinarios (MCO) arroja estimadores que sean Lineales, Insesgados, Eficientes y Consistentes, es decir:

I. Los estimadores ofrecen una solución lineal al sistema de ecuaciones.

II. El valor esperado del estimador es igual al estimador propio.

III. Los estimadores obtenidos presentan la menor varianza respecto de otros estimadores Lineales e Insesgados.

IV. Conforme se incrementa la muestra obtenida los estimadores tienden a sus valores "reales".

Para que los estimadores tengan dichas propiedades deben cumplirse los siguientes supuestos, Wooldridge (2013):

I. Linealidad de los parámetros: El modelo poblacional puede escribirse como

$$\boldsymbol{y = \beta_0 + \beta_1 x_1 + \beta_2 x_3 + \cdots + \beta_k x_k + e} \quad \text{ec. 5}$$

Es decir, debe poder representarse mediante una función lineal f$(x_1, x_2, x_3, \dots, x_k, e)$ con β_s igual a una constante para cada s de 0 a k.

II. Se toma una muestra aleatoria con n cantidad de observaciones de cada variable, suficiente y representativa para estimar los parámetros.

III. No existe una combinación lineal exacta entre las variables aleatorias y ninguna de las variables aleatorias es constante, recordar que en la representación vectorial X_0 es un vector de unos y no una variable aleatoria independiente.

IV. Se debe cumplir que:

$$\boldsymbol{E(e|x_1, x_2, x_3, \dots, x_k) = 0} \quad \text{ec. 6}$$

Dado que los errores no deben estar correlacionados a las variables independientes, el valor esperado del error, condicional en la información disponible de la muestra, debe ser cero. Este supuesto permite también establecer la insesgabilidad de los parámetros.

V. La varianza es constante, es decir:

$$\boldsymbol{Var(e|x_1, x_2, x_3, \dots, x_k) = \sigma^2} \quad \text{ec. 7}$$

Considerado como un supuesto secundario facilita la derivación de fórmulas para la varianza muestral, así como permitir que los estimadores mediante Mínimos Cuadrados Ordinarios (MCO) tengan la menor varianza entre todos los estimadores lineales e insesgados.

VI. Los errores se distribuyen normalmente:

$$\boldsymbol{e \sim Normal\ (0, \sigma^2)} \quad \text{ec. 8}$$

Este último supuesto es considerado para obtener las distribuciones muestrales exactas, utilizadas para las pruebas de hipótesis. Asimismo, este supuesto mejora la propiedad de eficiencia de los estimadores obtenidos por MCO.

Las variables también pueden estar indexadas en el tiempo, en dicho caso decimos que se tratan de series de tiempo. El índice t indica el tiempo de la variable, siendo t-1 un periodo anterior, t-2 dos periodos anteriores y así sucesivamente, lo mismo para los periodos posteriores. También es importante considerar que los periodos deben ser homogéneos, es decir que cada periodo está medido en las mismas unidades temporales.

Dado que las variables temporales tienen dirección y magnitud no es posible recombinar los datos contenidos en una población como una muestra aleatoria, en dicho caso se dice que se trata de un proceso estocástico y que toda la información disponible se conoce como realización de las variables aleatorias. Lo último es importante tanto para la obtención de los datos como para la estimación y pruebas de los modelos propuestos.

Las regresiones lineales que consideran a sus propias variables en estados, o tiempos, anteriores como variables explicativas, es decir, aquellas regresiones con x_t dependiente y x_{t-1}, x_{t-2}, …, x_{t-n}, como variables independientes son conocidos como Vectores Autorregresivos (VAR), útiles para analizar series de tiempo[5]. Bjørnland (2000) menciona que los modelos VAR son de fácil interpretación y disponibilidad, además de justificar que son capaces de identificar regularidades empíricas dejando entrever características que en los modelos de regresión lineal permanecen ocultos.

2.2.1. El modelo VAR y VEC

Los modelos VAR suelen tratar todas las variables observadas como endógenamente relacionadas, confiando más en técnicas estadísticas que en premisas teóricas específicas para imponer relaciones entre variables. Esto hace que los modelos VAR sean particularmente útiles para pronosticar y realizar análisis económicos a través de métodos como el análisis de respuesta al impulso o las descomposiciones de la varianza del error de pronóstico para explorar las interrelaciones entre variables, Lütkepohl (2007).

[5] Donde el subíndice indica el tiempo anterior de la variable x

Los modelos VAR (Vectores Autorregresivos) pueden verse como un modelo de ecuaciones simultáneas reducidas o como un caso particular de los modelos de series multivariadas. Su característica principal es que considera un número finito de periodos atrás en el tiempo de las variables de interés. Se puede denotar de la forma

$$\boldsymbol{y_t = \mu + \Gamma_1 y_{t-1} + \cdots + \Gamma_p y_{t-p} + \varepsilon_t} \qquad \text{ec. 9}$$

Aprovechando la notación de Greene (2002) así como algunas precisiones de Lütkepohl (2007) podemos definir los elementos de la ecuación de la forma siguiente. Y es un vector de K variables endógenas, μ es un vector con constantes de tamaño K, representado las intersecciones, Γ_i es una matriz de dimensión K × K, para i = 1, . . . , p, donde p es el número de rezagos que explican el modelo. Además, ε_t denota un proceso multivariado de ruido blanco, es decir, es un vector de perturbaciones no autocorrelacionadas que se distribuye N(0,Ω) , donde $\Omega = E(\varepsilon_t \varepsilon_t)$.

El Modelo de Corrección de Errores Vectoriales (VECM) es una representación utilizada en el análisis de modelos VAR (Autorregresión Vectorial), específicamente cuando las series de datos involucradas son integradas de orden 1, abreviado I(1), y potencialmente cointegradas. La fórmula del modelo VECM reparametriza la ecuación en niveles del VAR incorporando las diferencias de las variables con el fin de capturar la dinámica de corto plazo. Si las variables dentro del modelo están realmente cointegradas, esto implica que hay una o más relaciones estables a largo plazo entre ellas, y que vuelven al equilibrio después de una perturbación Lütkepohl (2005). El modelo VECM puede representarse del siguiente modo:

$$\boldsymbol{\Delta y_t = \Pi y_{t-1} + \Gamma_1 \Delta y_{t-1} + \cdots + \Gamma_{p-1} \Delta y_{t-p+1} + u_t} \qquad \text{ec. 10}$$

Donde $\Pi = \alpha\beta'$, siendo α y β de dimensión (K×r) y de rango r. La matriz β se denomina matriz de cointegración y α suele denominarse matriz de carga o coeficientes de ajuste. Los efectos de ajuste de largo plazo se observan en los coeficientes Γ_i.

2.3. Los modelos de inteligencia artificial

Las redes neuronales refieren a un método o tipo de aprendizaje que se encuentra dentro del gran rubro de la Inteligencia Artificial, el cual contiene al aprendizaje máquina y éste a su vez al aprendizaje profundo como subconjuntos. La pertenencia a cada conjunto depende del nivel de especialización en el que se ajuste la red neuronal, ofreciendo posibilidad de usar una red neuronal sin conocer con un nivel de detalle específico su estructura funcional.

La predicción mediante redes neuronales se debe a las ventajas que éstas presentan, tal y como describe Villalda et al. (2012):

> "Las redes neuronales pueden procesar información de forma más rápida que los computadores convencionales, pero tienen la desventaja de que no podemos seguir su respuesta paso a paso como se puede hacer al ejecutar un programa convencional en un ordenador por lo que no resulta fácil detectar los errores. Las redes neuronales artificiales son muy efectivas para resolver problemas complicados de clasificación y reconocimiento de patrones."

Un aspecto que no debe olvidarse sobre las redes neuronales artificiales es su capacidad de interpretar información con comportamiento no lineal, característica obtenida de asemejar a su contraparte biológica.

Ilustración 2-1: Representación de una neurona biológica

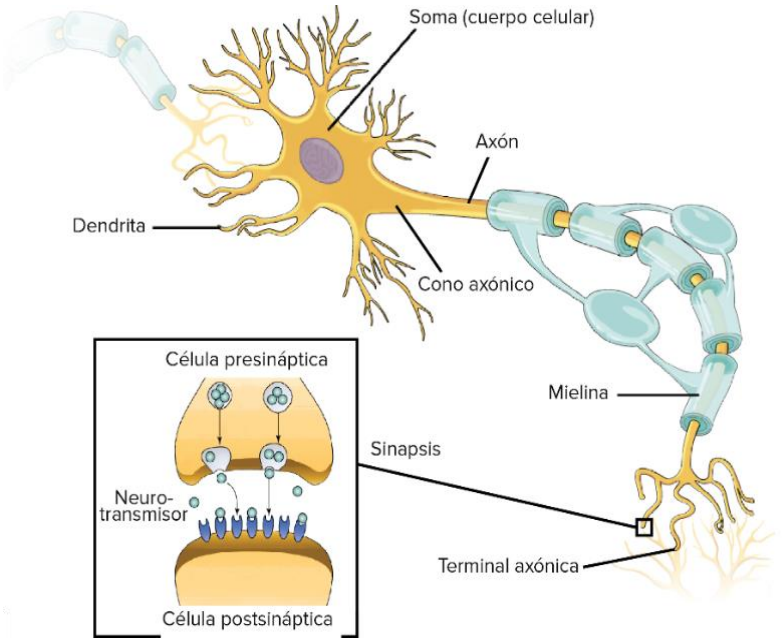


Imagen modificada de "Neuronas y células gliales" y "Sinapsis", de (Clark, Choi, y Douglas, 2020)

Una Red Neuronal Artificial se compone de muchas neuronas artificiales simples, y cada una de ellas se refiere a una unidad de procesamiento, nodo, o unidad lógica con límites[6], para fines prácticos se denominarán simplemente como neuronas. Así mismo, los valores sinápticos pueden ubicarse en diferentes niveles o capas, como si fungieran como filtros de otros valores sinápticos.

Diagrama 2-2: Esquema de una red neuronal computacional

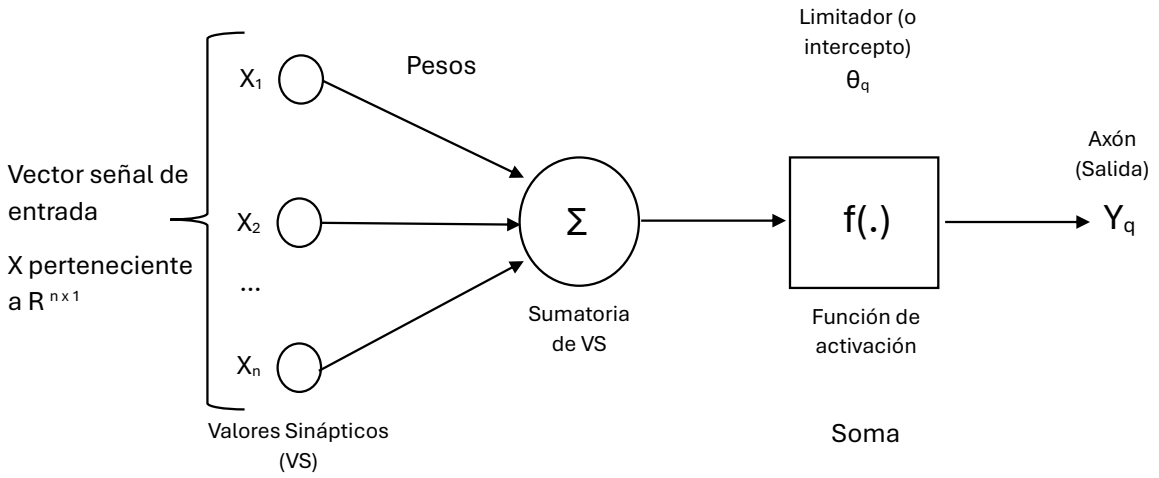


Imagen modificada de "Modelo no lineal de una neurona artificial" de Fredric y Kostanic (2001).

Tabla 2-1: Comparación de elementos de una neurona biológica y una neurona artificial

Neurona biológica	**Neurona artificial**
Estímulos	Vector de entrada
Neurotransmisores	Variables de X
Soma	Sumatoria de VS y función de activación
Axón	Salida
Dendritas	Pesos

Elaboración propia

[6] Los límites son los estados válidos que arrojan las funciones de activación.

Lo más destacable de una neurona artificial es que no existe una comparación exacta entre los elementos que la componen respecto a una neurona biológica, la primera busca asemejarse a la segunda en su parte funcional sin que ello implique estrictamente una relación uno a uno. En ese sentido, en otros textos podrán encontrarse diferencias respecto a las imágenes y tabla propuesta para entender sus similitudes.

De acuerdo con Fredric y Kostanic (2001) existen 4 componentes básicos de un modelo de redes neuronales:

- Una serie de sinapsis asociadas con valores transformados.
- Un mecanismo de suma que agrega todas las señales recibidas por el vector de señal de entrada, también conocido como nodo.
- Una función de activación que permite limitar la señal de salida de la neurona siempre y cuando la función sea no lineal. Generalmente los valores de salida asociados se acotan entre valores de [-1,1] o [0,1].
- Un valor de ajuste que disminuye el valor de suma de la entrada.

Las funciones de activación pueden ser o no lineales y dependen del tipo de clasificación que se busca para resolver un problema de optimización.

Tabla 2-2**:** Tipos de funciones de activación de una neurona artificial

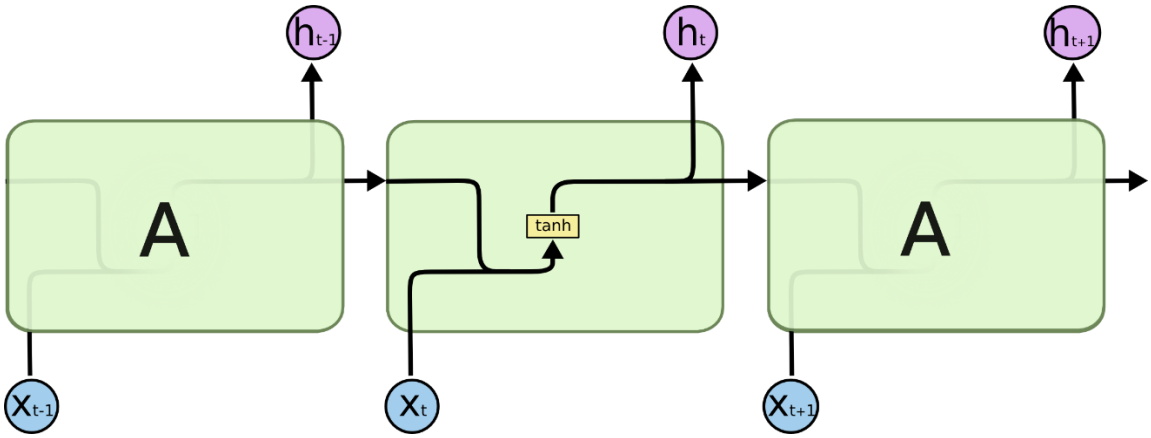
Funciones de activación lineales	**Funciones de activación no lineales**
• Función Identidad: Es una función simple que podría parecer trivial pero que puede resultar bastante útil. $y_q = f_i(v_q) = v_q$ • Limitador fuertemente simétrico (función signo): Es una función binaria con límites en 0 y 1 o bipolar con límites en -1 y 1. $f_{lfs}(v_q) = \begin{cases} -1, \wedge\ v_q < 0 \\ 0, \wedge\ v_q = 0 \\ 1, v_q > 0 \end{cases}$ $f_{lfs}(v_q) = \begin{cases} -1, \wedge\ v_q < 0 \\ 0, \wedge\ v_q = 0 \\ 1, v_q > 0 \end{cases}$ • Función lineal por partes: Igualmente puede tener límites binarios o bipolares y tener un valor lineal dentro de dichos rangos. $y_q = f_{lpp}(v_q) = \begin{cases} -1, \wedge\ v_q \leftarrow 1 \\ v_q, \wedge -1 \leq v_q \leq 1 \\ 1, v_q > 1 \end{cases}$	• Función sigmoide: La función más común, matemáticamente cuasicóncava y creciente con un rango binario. También es continua y diferenciable. $y_q = f_s(v_q) = \frac{1}{1 + e^{-\alpha v_q}}$ • Función hiperbólica tangente sigmoide (HTS): Es la variante con rango bipolar de la función sigmoide. $y_q = f_{hts}(v_q) = tanh(\alpha v_q)$ $y_q = \frac{1 - e^{-2\alpha v_q}}{1 + e^{-2\alpha v_q}}$

Elaboración propia con información de Fredric y Kostanic (2001)

2.3.1. Redes Neuronales recurrentes y la red LSTM

Uno de los tipos básicos de redes neuronales es la retroalimentación hacia adelante, también llamadas feed-forward, la cual se refiere a un tipo de arquitectura de red neuronal e+n la que las conexiones entre las unidades no forman un ciclo. Esto significa que en las redes neuronales feed-forward, la información fluye en una dirección, de entrada a salida, sin retorno. Esta arquitectura contrasta con las "redes neuronales recurrentes" (RNN), donde las conexiones entre las unidades pueden formar ciclos, lo que permite efectos de retroalimentación y memoria Sazli (2006).

Diagrama 2-3: Representación de una Red Neuronal Recurrente



Tomado de Gers et al. (2015)

Las redes neuronales recurrentes (RNN) se ocupan directamente de la naturaleza secuencial de series o datos indexados por el tiempo, este enfoque permite a las RNN utilizar los estados ocultos de una secuencia como una forma de memoria para procesar los pasos posteriores, capturando eficazmente las dependencias temporales. Dicha arquitectura les permite abarcar la secuencia, la continuidad y la memoria, lo que representa una desviación significativa del procesamiento simultáneo y sin estado característico de las redes tradicionales de feed-forward. Jurafsky y Martin (2023).

Las redes tipo Memoria de Largo-Corto plazo (LSTM por sus siglas en inglés) son un caso particular de las redes neuronales recurrentes diseñadas para abordar el problema del error de desaparición inherente a las RNN tradicionales. Este problema se refiere a la dificultad que tienen las RNN para aprender a unir intervalos de tiempo largos, donde el gradiente de la función de

pérdida se vuelve muy pequeño (desaparece), lo que dificulta el aprendizaje de dependencias a largo plazo. Las RNN tradicionales se enfrentan, en esa característica particular, al mismo problema que los estimadores lineales de modelos de Rezagos Distribuidos, en particular los modelos que incorporan el largo plazo son los de rezagos infinitos, en dichos modelos el impacto de los estimadores tiende a 0 conforme t tiende a infinito.

Sea y_t un proceso de rezagos infinitos

$$y_t = \alpha + \sum_{s=0}^{\infty} \beta_s x_{t-s} + e_t \qquad \text{ec. 11}$$

Para que sea posible la estimación de las betas debe asumirse un comportamiento específico, por ejemplo, un comportamiento geométrico

$$\beta_s = \beta \varphi^s \qquad \text{ec. 12}$$

Pero sabemos que si la serie es estacionaria entonces

$$\varphi < |\, 1 \,| \qquad \text{ec. 13}$$

$$\lim_{s \to \infty} \beta \varphi^s = 0 \qquad \text{ec. 14}$$

De este modo se muestra que estimar un modelo lineal que incorpora el largo plazo, por ejemplo, uno de Rezagos Distribuidos Infinitos, requiere de supuestos que limitan el comportamiento conforme los rezagos incrementan, en este caso con una tendencia estable a cero. De esta forma no se permite la posibilidad de perturbaciones significativas en el largo plazo o cambios en el comportamiento de los estimadores para un mismo modelo.

Las redes tipo LSTM abordan este problema mediante la introducción de celdas de memoria especiales que permiten un flujo de error constante a través de los estados internos de estas celdas, manteniendo efectivamente una memoria en secuencias largas. El componente central de LSTM que permite esta funcionalidad es el carrusel de errores constantes (CEC), que facilita la preservación del error a través del tiempo y el espacio, lo que permite aprender dependencias a largo plazo. La CEC trabaja en conjunto con varias puertas, Staudemeyer y Rothstein (2019):

1. Puerta de entrada: Esta puerta controla el flujo de la señal de entrada a la celda de memoria, determinando si la nueva entrada es lo suficientemente significativa como para recordarla.
2. Puerta de olvido (Forget Gate): Permite que la célula descarte información irrelevante, dejando espacio para nueva información y evitando así que la memoria se vea abrumada con datos irrelevantes.

3. Puerta de salida: Esta puerta controla el flujo de salida del estado de la celda al resto de la red, decidiendo si el estado actual es relevante para el problema en cuestión.

Diagrama 2-4: Funcionamiento de una neurona tipo LSTM

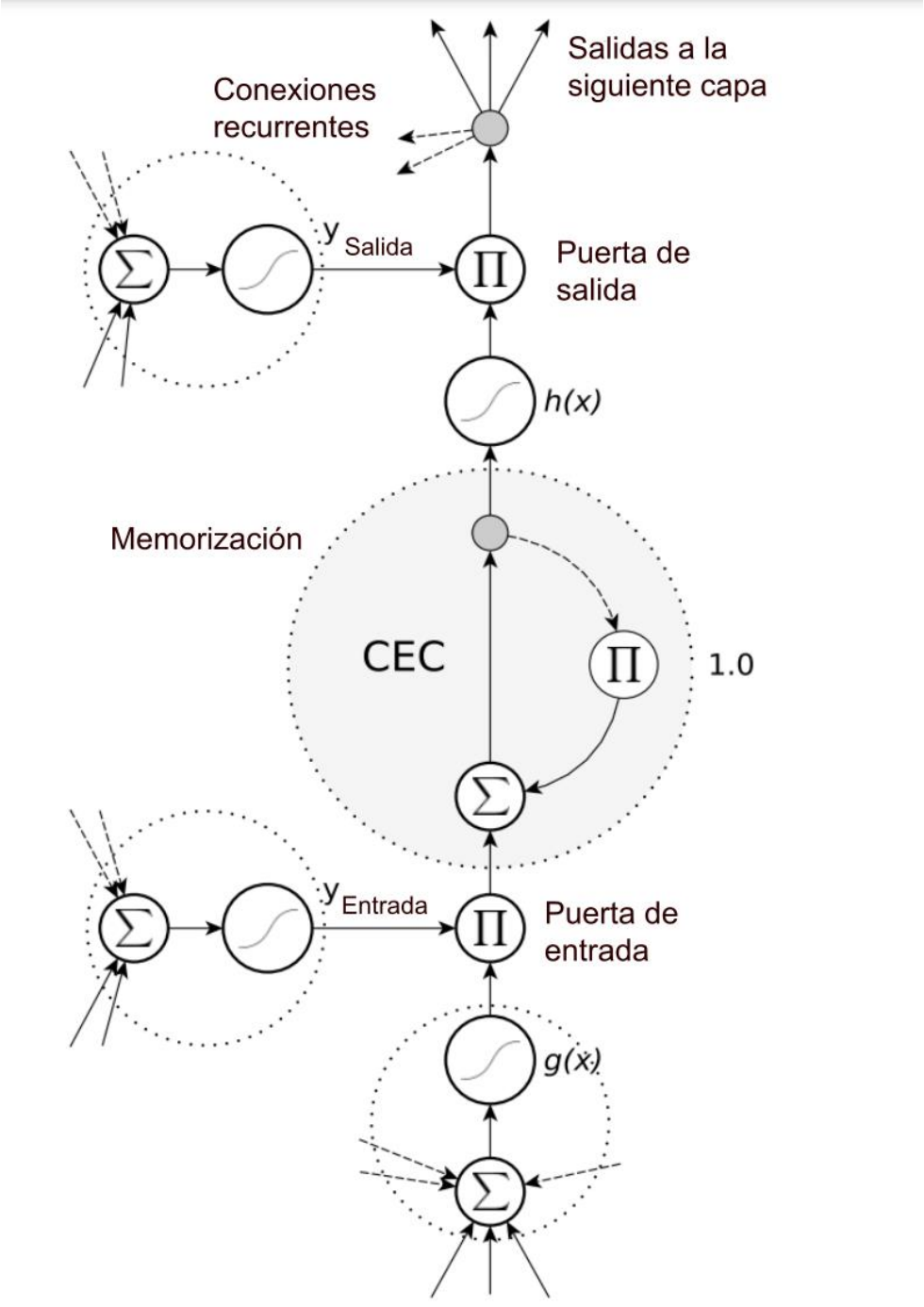


Ilustración modificada de Staudemeyer y Rothstein (2019):

El diagrama representa gráficamente el algoritmo de una neurona tipo LSTM, en ella se pueden observar las funciones de activación g(x), h(x), las sumatorias de valores sinápticos, las puertas y el mecanismo CEC con un peso unitario. Las puertas son esencialmente redes neuronales diferentes que aprenden qué almacenar, qué descartar y qué generar en función de la entrada que reciben. Su uso para series temporales ha demostrado un desempeño relevante en el campo económico y financiero, un ejemplo es su aplicación en los *stocks* o acciones del mercado de valores.

Al ser un tipo específico de una RNN las entradas y salidas de cada neurona tipo LSTM, mostrada en el diagrama 2-4, se conecta a otras neuronas del mismo tipo como se muestra en el diagrama 2-3, obteniendo las mismas propiedades predictivas en series de tiempo de las RNN y mejorando a través del funcionamiento interno de cada neurona, es decir, mediante funciones de activación adicionales y técnicas de memorización, en específico el método CEC.

Mehtab et al. (2009) usaron un modelo de regresión de aprendizaje profundo basado en LSTM (Long Short-Term Memory) y demostraron un rendimiento superior en las predicciones de índices bursátiles, específicamente para los valores del índice NIFTY 50, según el estudio. La investigación reveló que los modelos basados en LSTM tenían una capacidad significativamente mayor para extraer y aprender características de los datos de una serie temporal en comparación con los modelos predictivos basados en el aprendizaje automático, indicando así que los modelos LSTM, debido a su arquitectura, son capaces de capturar dependencias a largo plazo en datos de series temporales.

2.3.2. Los optimizadores de la red neuronal

Al igual que en los modelos econométricos el objetivo suele tratarse de un tema de optimización de una función, dado que los parámetros obtenidos siguen una serie de supuestos como linealidad, insesgabilidad, eficiencia y consistencia. En el caso de las redes neuronales se sigue un principio de optimización más general, Sun (2020) señala que una vez definido el problema, se debe elegir una función objetivo. Esta función, a menudo denominada función de pérdida o coste, cuantifica la distancia a la que se encuentran las predicciones del modelo de los resultados reales. El proceso de optimización tiene como objetivo minimizar esta función, lo que significa que busca reducir las diferencias entre los valores predichos y los verdaderos. Lo último puede generar un problema llamado sobreajuste, en el cuál la red se ajusta demasiado bien a los

datos de entrenamiento, pero es pobre en la predicción con nueva información o con la base de prueba.

Cuando existe una gran cantidad de información también debe tomarse en consideración la velocidad a la que la red neuronal "aprende", es decir, el tiempo que le toma acercarse a los valores más precisos. En el caso del presente trabajo dicha cuestión no representa un problema debido al tamaño de la base de datos.

Los algoritmos de optimización comunes incluyen el descenso de gradiente estocástico (SGD) y métodos adaptativos como Adam. Estos algoritmos ajustan iterativamente los parámetros en función del gradiente de la función objetivo con respecto a los parámetros, encontrando de esa manera el menor valor de pérdida. Dado que las funciones de optimización por regla general no son lineales es importante que el algoritmo no se atasque en mínimos locales, por ello es importante elegir el optimizador que mejor se adapte al modelo a predecir o estimar.

El algoritmo más utilizado, y del cual la praxis recomienda utilizar en primera instancia, es el de Adam. Este algoritmo está diseñado para la optimización de una función estocástica diferenciable. Sean los valores de entrada el conjunto de datos, una función objetivo C y un vector de parámetros iniciales θ_{init}, el algoritmo ADAM utiliza el siguiente procedimiento Liquet et al. (2024):

Tabla 2-3: Detalle de los pasos del algoritmo ADAM

Detalle del algoritmo	**Descripción**
1) $0 \rightarrow t$	Se inicia en 0 el contador para las iteraciones
2) $0 \rightarrow s$, $0 \rightarrow v$, $\theta_{init} \rightarrow \theta$	Se establecen los valores iniciales
3) Ciclo	
3.1) $\nabla C(\theta) \rightarrow g$	Se calculan los gradientes
3.2) $\beta v + (1 - \beta) g \rightarrow v$	Se actualiza el valor del primer momento
3.3) $\gamma s + (1 - \gamma) g \rightarrow s$	Se actualiza el valor del segundo momento
3.4) $\frac{v}{1-\beta^{t+1}} \rightarrow \tilde{v}$	Corrección de sesgo del primer momento
3.5) $\frac{s}{1-\beta^{t+1}} \rightarrow \check{s}$	Corrección de sesgo del segundo momento
3.6) $\theta - \alpha \frac{1}{\sqrt{\check{s}}+\varepsilon} \tilde{v} \rightarrow \theta$	Actualización de parámetros
g) $t + 1 \rightarrow t$	Actualización del ciclo de iteración

4) El paso 3 se repite hasta que la condición de término se satisface

Elaboración propia con datos de Liquet et al. (2024)

La tabla requiere de un conjunto de precisiones, el bucle continúa hasta que los parámetros θ convergen, de modo que la salida del algoritmo Adam son los parámetros θ optimizados, los valores de β y γ son las tasas de decaimiento exponencial para el primer y segundo momento respectivamente. Dado que los valores iniciales de los dos primeros momentos se establecen en cero existe por tanto un sesgo que busca corregirse durante las iteraciones, así mismo ε es una constante pequeña que evita la división por cero y α es la tasa de aprendizaje. Las constantes α, β y γ se encuentran dentro del intervalo [0,1] y los valores por defecto en el software son $\alpha = 0.001$, $\beta = 0.9$, $\gamma = 0.999$, $\varepsilon = 1e\text{-}07$.

Capítulo 3. Los modelos y sus parámetros

3.1. Recopilación de la información y limpieza de datos

Lo que la literatura sugiere es que además del capital, la fuerza de trabajo y la tecnología existe un componente que no se ha integrado en los modelos ortodoxos para la estimación del PIB, esto es la inflación. En ese sentido se esperaría que exista una relación de cointegración de la forma:

$$PIB_t = \beta_0 + \beta_1 U_t + \beta_2 K_t + \beta_3 INF_t + e_t \qquad \text{ec. 15}$$

En donde β_0 correspondería al residual de Solow o el nivel de tecnología del país a analizar, U la tasa de desempleo, K el nivel de acervo de capital disponible, INF la tasa de inflación, y t es el periodo de tiempo en intervalos trimestrales. Debido a que el PIB y el acervo de capital en niveles tiene órdenes de magnitud completamente distintos a la tasa de desempleo y la tasa de inflación, se realiza el ajuste por logaritmos, quedando la relación de la siguiente manera:

$$\ln(PIB_t) = \beta_0 + \beta_1 U_t + \beta_2 \ln(K_t) + \beta_3 INF_t + e_t \qquad \text{ec. 16}$$

De forma que los resultados pueden ser interpretados de una forma más intuitiva. No se utiliza la tasa de crecimiento del PIB porque se busca obtener un valor potencial que no sería calculable mediante el PIB en diferencias, así mismo no se utiliza la tasa de crecimiento del acervo de capital debido a que dicha diferencia es equivalente al flujo de capital, el cual se interpreta como un factor de demanda y el objeto de análisis del presente trabajo busca encontrar la relación desde el lado de la oferta.

Los datos para el PIB en niveles se obtienen de la página oficial del Instituto Nacional de Estadística e Informática (INEGI), se descargan las series originales con base 2018 para incorporar los efectos estacionales en el modelo. Por su parte la tasa de inflación se obtiene del Índice de Nacional de Precios al Consumidor (INPC) como la tasa de crecimiento del INPC respecto a su mismo mes del año anterior.

Debido a que no existe una serie oficial del Acervo de Capital para México de forma trimestral, se utiliza el método de inventarios perpetuos con ajuste de Almon para obtener la serie. Gutierrez y Moreno (2022) resumen el ajuste de Almon para la normalización del Acervo del capital mediante la fórmula:

$$Faj_t = (1 - \delta) * Faj_{t-1} + 1 \qquad \text{ec. 17}$$

Donde Faj es el factor de ajuste; δ la tasa de depreciación y t el periodo, los autores asumen que el primer dato sea igual a 1. Posteriormente se agrega el factor de ajuste a la serie de capital

$$SCA_t = \frac{SC_t / Faj_t}{\delta} \qquad \text{ec. 18}$$

Donde SCA es el acervo de capital ajustado y SC el acervo de capital. Para la construcción del Acervo de Capital sin ajuste se integraron los valores de la serie de Formación Bruta de Capital Fijo, base 2018, que publica el INEGI de forma desagregada entre *Bienes inmuebles* y *Maquinaria, equipo y otros activos* fijos mediante la fórmula:

$$SC_t = (1 - \delta) * SC_{t-1} + FBKF_t \qquad \text{ec. 19}$$

Donde FBKF es la Formación Bruta de Capital Fijo desagregada, en ese sentido, cabe destacar que se estimaron dos tasas de depreciación para las series, dicha estimación se obtuvo mediante iteración, buscando el valor de depreciación que hiciera que los flujos acumulados y ajustados con la corrección de Almon fuesen equivalentes en el periodo base (2018) a los publicados por INEGI. El motivo de realizar la estimación de las tasas de depreciación de forma desagregada es porque los Bienes inmuebles tienen una tasa de depreciación distinta y no ponderable a la tasa de Maquinaria, equipo y otros activos. Los valores de δ fueron 0.039116 y 0.140158 respectivamente, lo cual equivale a una pérdida del valor del capital en 25.64 y 7.13 años respectivamente. Las estimaciones de las tasas de depreciación son similares a las tasas de depreciación obtenidas con otros métodos[7].

La serie de tasa de desempleo se obtuvo para el periodo 1993-2000 de la Encuesta Nacional de Empleo Urbano (ENEU) y para los años posteriores de la Encuesta Nacional de Ocupación y Empleo (ENOE). Dado que la ENEU no reporta una serie de la tasa de desempleo, esta se obtuvo mediante una estimación con los microdatos históricos por cuatrimestre.

A pesar de que las series de empleo tienen algunas diferencias en la metodología usada por el INEGI la serie completa es muy similar a otras series oficiales, como la serie de *Desempleo, total (% de la fuerza laboral total) (estimación modelada de la OIT)–México* reportada por el Banco Mundial en su página web, la cual a su vez es recuperada de una estimación de la Organización Internacional del Trabajo.

[7] Para más detalle de otros métodos y tasas obtenidas se pueden revisar los trabajos de Erumban (2008), Fisher et al. (2005), Bokhari et al. (2016) por citar algunos.

3.2. Aplicación del modelo VAR y VEC

Siguiendo la relación de cointegración esperada se plantea el modelo VAR:

$$y_t = \mu + \Gamma_1 y_{t-1} + \cdots + \Gamma_p y_{t-p} + \varepsilon_t \quad \text{ec. 20}$$

$$y = [\ln(\text{PIB}), \text{U}, \ln(\text{K}), \text{INF}]^{\text{T}} \quad \text{ec. 21}$$

Es decir, y es el vector de las variables de interés propuestas al inicio de la presente sección. Por otra parte, p es un orden de rezagos a obtener mediante el criterio de máxima verosimilitud. Si $L_n(k)$ es la función de verosimilitud de un modelo VAR(p) gaussiano, entonces debemos usar estos criterios para determinar el orden p del VAR:

$$\hat{\text{p}} = \text{argmin}_p c_n^{VAR}(p) \quad \text{ec. 22}$$

Donde:

Akaike: $$c_n^{VAR}(p) = \ln\left(det\left(\Sigma_p\right)\right) + 2\ (m + m^2 p)/n \quad \text{ec. 23}$$

Hannan-Quinn: $$c_n^{VAR}(p) = \ln\left(det\left(\Sigma_p\right)\right) + 2\ (m + m^2 p)\ln\ (\ln(n))/n \quad \text{ec. 24}$$

Schwarz: $$c_n^{VAR}(p) = \ln\left(det\left(\Sigma_p\right)\right) + 2\ (m + m^2 p)\ln\ (n)/n \quad \text{ec. 25}$$

Para n grados de libertad de las observaciones temporales (observaciones totales menos número de rezagos) y m variables Bierens (2004). Toda vez obtenido el orden de rezagos óptimo se procede a revisar la normalidad, autocorrelación y heteroscedasticidad de los residuos, dado que el supuesto para que exista una relación de cointegración es que los errores sean estacionarios, es decir, sea $\ln(\text{PIB}_t) \sim \text{I}(1)$, $\text{U}_t \sim \text{I}(1)$, $\ln(\text{K}_t) \sim \text{I}(1)$ e $\text{INF}_t \sim \text{I}(1)$, entonces $\varepsilon_t \sim \text{I}(0)$. "Aunque la normalidad no es una condición necesaria para la validez de muchos de los procedimientos estadísticos relacionados con los modelos VAR, las desviaciones de la hipótesis de normalidad pueden indicar que es posible mejorar el modelo" Lütkepohl (2007).

La importancia de comprobar si hay una interrupción en un conjunto de datos, especialmente cuando se examina la posibilidad de una relación de cointegración a largo plazo, es crucial puesto que las pruebas tradicionales, como las pruebas convencionales de Dickey-Fuller aumentado (ADF), podrían no rechazar la hipótesis nula de la ausencia de cointegración, lo que podría llevar a los investigadores a concluir incorrectamente que no existe una relación a largo plazo entre las variables. Al tener en cuenta las rupturas o los cambios de régimen en el modelo, el análisis podría detectar una cointegración que no era evidente sin tener en cuenta dichos cambios estructurales Gregory y Hansen (1996).

Una vez que las pruebas arrojen evidencia de una relación de cointegración se procede a revisar la relación de corto plazo de las variables, no debe olvidarse que una relación de cointegración implica que las variables tienen un equilibrio en el largo plazo. Para ello se utiliza el Vector de Corrección de Error, para revisar el comportamiento de ajuste en el corto plazo, el modelo a utilizar es:

$$\Delta y_t = \Pi y_{t-1} + \Gamma_1 \Delta y_{t-1} + \cdots + \Gamma_{p-1} \Delta y_{t-p+1} + u_t \qquad \text{ec. 26}$$

$$\Delta y = [\Delta \ln(\text{PIB}), \Delta \text{U}, \Delta \ln(\text{K}), \Delta \text{INF}]^{\text{T}} \qquad \text{ec. 27}$$

Sea $\Delta\ln(\text{PIB}_t) \sim \text{I}(0)$, $\Delta\text{U}_t \sim \text{I}(0)$, $\Delta\ln(\text{K}_t) \sim \text{I}(0)$ y $\Delta\text{INF}_t \sim \text{I}(0)$, entonces $u_t \sim \text{I}(0)$. Para identificar la validez del modelo, al igual que en el modelo VAR se revisa la normalidad, autocorrelación y heteroscedasticidad.

3.3. Aplicación del modelo de Inteligencia Artificial

Para la creación del modelo de IA es importante recurrir a la selección de paradigmas en redes neuronales, incluyendo la determinación de capas ocultas, neuronas y funciones de transferencia. Los criterios de evaluación para medir el rendimiento del modelo en conjuntos de pruebas y validación son fundamentales en este proceso de entrenamiento, el cual implica decisiones sobre iteraciones, velocidad de aprendizaje e impulso.

Las capas ocultas permiten a la red aprender patrones y relaciones complejos en los datos, mejorando así la potencia computacional de la red y su capacidad para generalizar a partir de los datos de entrada para hacer predicciones o decisiones precisas. La implementación requiere atención a transformaciones de datos, escalado y coherencia de parámetros y la metodología de entrenamiento-prueba requiere de la aleatorización como una técnica de regularización para evitar el sobreajuste, el cual es un serio problema que se deriva del número de neuronas de capa oculta utilizados.

Debido a que no se pueden realizar pruebas controladas para experimentar con los modelos económicos los pronósticos se evalúan contra información que ya ha sucedido, de este modo los datos se dividen en un conjunto de entrenamiento y un conjunto de evaluación.

En ese orden de ideas y utilizando la información detallada en el capítulo anterior se plantea el uso de una red neuronal con 1 capa, 30 neuronas tipo LSTM, 100 iteraciones, un lote de 32 valores por conjunto, el optimizador Adam y una función de pérdida por mínimos errores

cuadrados. Para la selección del número de rezagos deben realizarse estimaciones con cada uno de ellos y seleccionar el que tenga menor diferencia entre los errores del conjunto de entrenamiento y el de prueba.

$$PEE_j = \frac{\sum_{i=1}^{n} | \ln(PIB_i)_{pred_{train}} - \ln(PIB_i)_{real} |}{n} \quad \text{ec. 28}$$

$$PEP_j = \frac{\sum_{i=1}^{n} | \ln(PIB_i)_{pred_test} - \ln(PIB_i)_{real} |}{n} \quad \text{ec. 29}$$

$$ME = \text{Min} \{ | PEP_j - PEE_j | \} \quad \text{ec. 30}$$

Donde PEE es el Promedio de Errores de Entrenamiento, pred_train es el subíndice del conjunto de valores de entrenamiento, pred_test el subíndice del conjunto de valores de prueba, n es el número de observaciones o periodos de cada set, PEP es el Promedio de Errores de Prueba, ME es la Mejor Estimación y j son los números de rezagos utilizados.

La mejor estimación parte del supuesto y la evidencia empírica sobre la tendencia al sobreajuste por parte de las redes neuronales artificiales y por tanto el criterio principal para seleccionar la mejor configuración de estas es: aquella que mejor se ajusta en el conjunto de prueba dado el ajuste en el conjunto de entrenamiento. En otras palabras, toda red tiende a cero en el error del conjunto de entrenamiento, pero si al momento de predecir se "equivoca" demasiado (en el conjunto de prueba) la red no es útil.

Cabe destacar que las redes con neuronas LSTM utilizan dos criterios de filtrado de información, el lote y el número de rezagos. El lote se utiliza para la "memoria" de largo plazo y el número de rezagos para la de corto plazo.

Para la replicación de los resultados expuestos en el presente trabajo se utiliza una semilla específica, en los modelos generales de redes neuronales dicha semilla o valor inicial es pseudo-aleatorio.

Capítulo 4. Resultados

La estimación de la prueba Dickey-Fuller aumentada se realiza para las variables lpib, u, lakta e inf_an, donde: lpib representa el logaritmo del PIB trimestral, u es la tasa de desempleo trimestral, lakta es el logaritmo del acervo de capital trimestral con corrección de Almon e inf_an es la inflación anual por trimestre.

Tabla 4-1: Resultados pruebas Dickey-Fuller aumentadas para las variables en niveles

Variable	**Valor z muestral**	**Aproximación del valor p-value de MacKinnon**
lpib	-1.604	0.4817
u	-2.126	0.2343
inf_an	-1.733	0.4144
lakta	-1.733	0.7298
Hipótesis nula: La serie contiene al menos una raíz unitaria. No se rechazan las hipótesis nulas de raíz unitaria, las series en niveles son no estacionarias.		

Elaboración propia

Tabla 4-2: Resultados pruebas Dickey-Fuller aumentadas para las variables en primeras diferencias

Variable	**Valor z muestral**	**Aproximación del valor p-value de MacKinnon**
Δlpib	-15.285	0
Δu	-13.057	0
Δinf_an	-5.948	0
Δlakta	-6.266	0
Hipótesis nula: La serie contiene al menos una raíz unitaria. No se rechazan las hipótesis nulas de raíz unitaria, las series en primeras diferencias son estacionarias.		

Elaboración propia

Las pruebas de Dickey-Fuller muestran que todas las variables obtenidas en niveles son no estacionarias e integradas de orden 1, por lo cual se procede a realizar pruebas de cointegración. La prueba de Gregory-Hansen muestra evidencia de un corte estructural en el primer trimestre del 2018.

Tabla 4-3: Prueba de cointegración de Gregory-Hansen en modelo 3, cambio de nivel con tendencia

Procedimiento ADF	
Estadístico t	-6.753441
Retraso	0
Ruptura	2018T1
Procedimiento de Phillips	
Estadístico Za	-64.8911
Ruptura de Za	2018Q2
Estadístico Zt	-6.736508
Ruptura de Zt	2018Q1
Valores de rechazo para Gregory-Hansen	
ADF	-5.5
Phillips (t)	-5.5
Phillips alpha	-58.33

Elaboración propia

Tabla 4-4: Sistema VAR, máximo orden de retardos 9

Retardos	log.veros	p(RV)	AIC*	BIC*	HQC*
1	1213.476	-	-25.249	-23.824	-24.674
2	1292.552	0	-26.621	-24.757	-25.868
3	1338.924	0	-27.281	-24.978	-26.352
4	1410.58	0	-28.491	-25.75	-27.385
5	1473.628	0	-29.514	-26.334	-28.23
6	1521.437	0	-30.205	-26.586*	-28.745
7	1547.659	0	-30.427	-26.371	-28.79
8	1579.611	0	-30.774	-26.279	-28.96
9	1621.96	0	-31.346*	-26.413	-29.355*

*Los asteriscos de abajo indican los mejores (es decir, los mínimos), valores de cada criterio de información, AIC = criterio de Akaike, BIC = criterio bayesiano de Schwarz y HQC = criterio de Hannan-Quinn.

Elaboración propia

La estimación del VAR se hace con el vector de variables endógenas Y = (lpib, u, lakta, inf_an). Para la selección de número de rezagos se utiliza el criterio de Akaike (AIC) y Hannan-Quinn (HQC). Ambos coinciden que el mejor criterio es el de mayor número de rezagos cuando se incluyen las variables dummies para corrección de normalidad, sin embargo, los datos muestran autocorrelación de residuales en diferentes niveles dependiendo el número de rezagos seleccionados. Tras varias pruebas con diferentes órdenes de rezagos, que cumplen los criterios AIC y HQC, se encuentra que con 9 rezagos es suficiente para atacar el problema de autocorrelación. Los resultados del VAR se desglosan en el Apéndice A.

Tabla 4-5: Prueba de Causalidad de Granger para VAR / Prueba de Wald de exogeneidad

Variable dependiente: lpib			
Excluida	*Chi-2*	*GL*	*Prob.*
u	12.84249	9	0.1699
lakta	60.53686	9	0.0000
inf_an	26.73459	9	0.0015
Todas	164.4814	27	0.0000
Hipótesis nula: Las variables excluidas no causan en términos de Granger a lpib			

Elaboración propia

Tabla 4-6: Pruebas de normalidad residual VAR con ortogonalización Cholesky (Lutkepohl)

Componente	*Jarque-Bera*	*gl*	*Prob.*
1	2.2473	2	0.3251
2	3.0870	2	0.2136
3	1.2512	2	0.5349
4	1.1978	2	0.5494
Conjunta	7.7833	8	0.4549
Hipótesis nula: Los residuos son normales multivariados			

Elaboración propia

Tabla 4-7: Pruebas de normalidad residual VAR con ortogonalización de correlación residual (Doornik-Hansen)

Componente	Jarque-Bera	gl	Prob.
1	2.8835	2	0.2365
2	3.9720	2	0.1372
3	0.2693	2	0.874
4	3.1320	2	0.2089
Conjunta	10.2570	8	0.2475
Hipótesis nula: Los residuos son normales multivariados			

Elaboración propia

Tabla 4-8: Pruebas de heterocedasticidad residual VAR (niveles y cuadrados)

Prueba articular:		
Chi-2	*gl*	*Prob.*
790.6412	800	0.5864

Elaboración propia

Tabla 4-9: Pruebas de heterocedasticidad residual VAR por componentes individuales

Dependiente	*R-2*	*F(80,11)*	*Prob.*	*Chi-2(80)*	*Prob.*
res1*res1	0.8584	0.8334	0.6992	78.9706	0.5115
res2*res2	0.8705	0.9241	0.6138	80.0843	0.4763
res3*res3	0.8852	1.0604	0.4955	81.4400	0.4341
res4*res4	0.8607	0.8494	0.6840	79.1819	0.5048
res2*res1	0.8824	1.0320	0.5188	81.1836	0.4420
res3*res1	0.8177	0.6169	0.8914	75.2324	0.6299
res3*res2	0.8913	1.1275	0.4435	81.9996	0.4171
res4*res1	0.8905	1.1185	0.4503	81.9280	0.4192
res4*res2	0.9120	1.4247	0.2669	83.9023	0.3609
res4*res3	0.8994	1.2298	0.3731	82.7481	0.3946

Elaboración propia

Tabla 4-10: Pruebas de LM de correlación serial residual VAR por rezago

Retardo	*LRE* estad.*	*gl*	*Prob.*	*F-estad.*	*gl*	*Prob.*
1	21.26935	16	0.1684	1.366540	(16, 122.8)	0.1695
2	14.87855	16	0.5336	0.932165	(16, 122.8)	0.5348
3	27.18909	16	0.0394	1.788426	(16, 122.8)	0.0399
4	8.406064	16	0.9358	0.513497	(16, 122.8)	0.9361
5	26.70995	16	0.0448	1.753559	(16, 122.8)	0.0453
6	13.96849	16	0.6011	0.872030	(16, 122.8)	0.6022
7	15.01144	16	0.5238	0.940981	(16, 122.8)	0.5251
8	26.31183	16	0.0498	1.724685	(16, 122.8)	0.0503
9	14.74132	16	0.5437	0.923070	(16, 122.8)	0.5449
10	14.70574	16	0.5463	0.920713	(16, 122.8)	0.5475
Hipótesis nula: No hay correlación serial con rezagos hasta h						

Elaboración propia

Tabla 4-11: Pruebas de LM de correlación serial residual VAR por rezago acumulado

Retardo	*LRE* estad.*	*gl*	*Prob.*	*F-estad.*	*gl*	*Prob.*
1	21.26935	16	0.1684	1.366540	(16, 122.8)	0.1695
2	32.44047	32	0.4451	1.018568	(32, 134.4)	0.4510
3	57.00387	48	0.1751	1.223518	(48, 125.3)	0.1879
4	66.95944	64	0.3758	1.045069	(64, 111.9)	0.4134
5	82.06102	80	0.4152	1.004372	(80, 97.1)	0.4891
6	98.30444	96	0.4156	0.975750	(96, 81.7)	0.5481
7	125.5798	112	0.1795	1.074619	(112, 66.1)	0.3793
8	146.4621	128	0.1263	1.038792	(128, 50.4)	0.4496
9	193.5364	144	0.0037	1.261088	(144, 34.5)	0.2155
10	231.9340	160	0.0002	1.137019	(160, 18.7)	0.3930
Hipótesis nula: No hay correlación serial con rezagos de 1 a h						

Elaboración propia

Las pruebas de causalidad de Granger en el VAR con 9 rezagos muestran una significancia robusta para rechazar la no causalidad de Granger. Con las variables dummies exógenas, correspondientes a diferentes choques económicos, se normalizan los errores y se pasan satisfactoriamente las pruebas de no heterocedasticidad, y no autocorrelación.

Dichos resultados muestran una posible relación de cointegración que se evalúa mediante la prueba de Johansen (1987) y mediante el modelo VEC correspondiente.

Tabla 4-12: Prueba de Cointegración de Johansen

Posibles No. de EC(s)	*Eigenvalor*	*Estadístico (Max-Eigen)*	*0.05*	*Prob.***
			Valor crítico	*Valor crítico*
Ninguno*	0.304814	54.85494	47.85613	0.0096
A lo sumo 1	0.12494	21.40591	29.79707	0.3328
A lo sumo 2	0.083357	9.127321	15.49471	0.3537
A lo sumo 3	0.012099	1.119923	3.841465	0.2899
La prueba de traza indica 1 ecuación(es) de cointegración en el nivel 0,05				
* denota rechazo de la hipótesis en el nivel de 0,05				
**Valores p de MacKinnon-Haug-Michelis (1999)				

Elaboración propia

Tabla 4-13: Prueba de rango de cointegración sin restricciones (valor propio máximo)

Posibles No. de EC(s)	*Eigenvalor*	*Estadístico (Max-Eigen)*	*0.05*	*Prob.***
			Valor crítico	*Valor crítico*
Ninguno*	0.304814	33.44904	27.58434	0.0078
A lo sumo 1	0.12494	12.27858	21.13162	0.5204
A lo sumo 2	0.083357	8.007399	14.2646	0.378
A lo sumo 3	0.012099	1.119923	3.841465	0.2899
La prueba de valores propios máximos indica 1 ecuación de cointegración en el nivel 0,05				

* denota rechazo de la hipótesis en el nivel de 0,05
**Valores p de MacKinnon-Haug-Michelis (1999

Elaboración propia

Los resultados arrojan 1 vector de cointegración al 5% de significancia. Siendo la ecuación de cointegración del lpib normalizado la siguiente:

$$\Delta lpib_t = 2.175670 + 0.2536(lpib_{t-1} + 5.066\, u_{t-1} - 1.402\, lakta_{t-1} - 0.597\ \text{inf_}an_{t-1}) + \Gamma_1 \Delta y_{t-1} + \cdots + \Gamma_{t-8} \Delta y_{t-8} + \gamma \Delta D_t + u_t \quad \text{ec. 31}$$

Donde los coeficientes Γ de las 8 diferencias de las variables se listan en el Apéndice B, también se agregan los coeficientes γ para las variables dummies correspondientes a la normalización de los errores por impactos económicos inusuales, como la crisis del peso en 1994, el boom de las ".com" y la crisis del sector hipotecario del 2008.

Se revisa la significancia de los coeficientes obtenidos, tanto de los asociados al vector de corrección de errores como al vector de cointegración.

Tabla 4-14: Pruebas de significancia de los parámetros de corrección y vector de cointegración

Restricciones de cointegración:	**Hipótesis nula**	**Iteraciones para convergencia**	**Chi-cuadrado (1)**	**Probabilidad**
a(1,1)=0	El parámetro de corrección de error asociado al PIB no es significativo en el modelo.	11	7.1533	0.0075
a(2,1)=0	El parámetro de corrección de error asociado a la tasa de desempleo no es significativo en el modelo.	7	3.0283	0.0818

a(3,1)=0	El parámetro de corrección de error asociado al Acervo de Capital Total no es significativo en el modelo.	32	20.9783	0.0000
a(4,1)=0	El parámetro de corrección de error asociado a la inflación anual no es significativo en el modelo.	5	0.4397	0.5073
b(1,1)=0	El parámetro del vector de cointegración asociado al PIB no es significativo en el modelo.	33	20.8844	0.0000
b(1,2)=0	El parámetro del vector de cointegración asociado a la tasa de desempleo no es significativo en el modelo.	23	20.0808	0.0000
b(1,3)=0	El parámetro del vector de cointegración asociado al Acervo de Capital Total no es significativo en el modelo.	53	21.1110	0.0000
b(1,4)=0	El parámetro del vector de cointegración asociado a la inflación anual no es significativo en el modelo.	7	9.0246	0.0027

Elaboración propia

Se realizan las pruebas de no autocorrelación, no heterocedasticidad y normalidad del modelo VEC.

Tabla 4-15: Pruebas de normalidad residual VEC con ortogonalización Cholesky (Lutkepohl)

Componente	*Jarque-Bera*	*gl*	*Prob.*
1	0.8876	2	0.6416
2	1.1811	2	0.5540
3	1.0669	2	0.5866
4	3.4107	2	0.1817
Conjunta	6.5464	8	0.5863
Hipótesis nula: Los residuos son normales multivariados			

Elaboración propia

Tabla 4-16: Pruebas de normalidad residual VEC con ortogonalización de correlación residual (Doornik-Hansen)

Componente	*Jarque-Bera*	*gl*	*Prob.*
1	2.4075	2	0.3001
2	1.3337	2	0.5133
3	0.6681	2	0.7160
4	3.2208	2	0.1998
Conjunta	7.6303	8	0.4704
Hipótesis nula: Los residuos son normales multivariados			

Elaboración propia

Tabla 4-17: Pruebas de heterocedasticidad residual VEC (niveles y cuadrados)

Prueba conjunta:		
Chi-2	*df*	Prob.
760.3631	770	0.5908

Elaboración propia

Tabla 4-18: Pruebas de heterocedasticidad residual VEC por componentes individuales:

Dependiente	*R-2*	*F(77,14)*	*Prob.*	*Chi-2(77)*	*Prob.*
res1*res1	0.892194	1.504719	0.1988	82.08188	0.3248
res2*res2	0.876527	1.290710	0.3075	80.64044	0.3660
res3*res3	0.801547	0.734359	0.8072	73.74232	0.5841
res4*res4	0.792013	0.692364	0.8463	72.86524	0.6123
res2*res1	0.896718	1.578586	0.1710	82.49805	0.3134
res3*res1	0.842010	0.969005	0.5681	77.46495	0.4637
res3*res2	0.780690	0.647228	0.8847	71.82347	0.6454
res4*res1	0.875181	1.274841	0.3174	80.51669	0.3697
res4*res2	0.825197	0.858315	0.6811	75.91814	0.5135
res4*res3	0.751032	0.548470	0.9501	69.09498	0.7277

Elaboración propia

Tabla 4-19: Pruebas de LM de correlación serial residual VAR por rezago

Retardo	*LRE* estad.*	*gl*	*Prob.*	*F-estad.*	*gl*	*Prob.*
1	17.94606	16	0.3271	1.136341	(16, 138.1)	0.3281
2	10.26450	16	0.8525	0.632698	(16, 138.1)	0.8529
3	20.82238	16	0.1854	1.331894	(16, 138.1)	0.1863
4	20.84822	16	0.1844	1.333668	(16, 138.1)	0.1853
5	21.86182	16	0.1477	1.403521	(16, 138.1)	0.1485
6	21.53732	16	0.1588	1.381105	(16, 138.1)	0.1596
7	9.379790	16	0.8970	0.576385	(16, 138.1)	0.8973
8	23.44716	16	0.1023	1.513760	(16, 138.1)	0.1030
9	14.75691	16	0.5425	0.924002	(16, 138.1)	0.5435
10	8.441804	16	0.9346	0.517056	(16, 138.1)	0.9348
Hipótesis nula: No hay correlación serial con rezagos hasta h						

Elaboración propia

Tabla 4-20: Pruebas de LM de correlación serial residual VAR por rezago acumulado

Retardo	*LRE* estad.*	*gl*	*Prob.*	*F-estad.*	*gl*	*Prob.*
1	17.94606	16	0.3271	1.136341	(16, 138.1)	0.3281
2	30.65755	32	0.5345	0.957186	(32, 152.8)	0.5389
3	47.27037	48	0.5027	0.981577	(48, 144.6)	0.5156
4	61.44709	64	0.5673	0.944248	(64, 131.5)	0.5947
5	82.34434	80	0.4067	1.016794	(80, 116.8)	0.4626
6	102.0586	96	0.3169	1.044918	(96, 101.6)	0.4131
7	124.5810	112	0.1962	1.089950	(112, 86.0)	0.3395
8	133.9630	128	0.3414	0.951089	(128, 70.3)	0.6023
9	160.1336	144	0.1695	0.977095	(144, 54.5)	0.5544
Hipótesis nula: No hay correlación serial en los rezagos 1 a h						

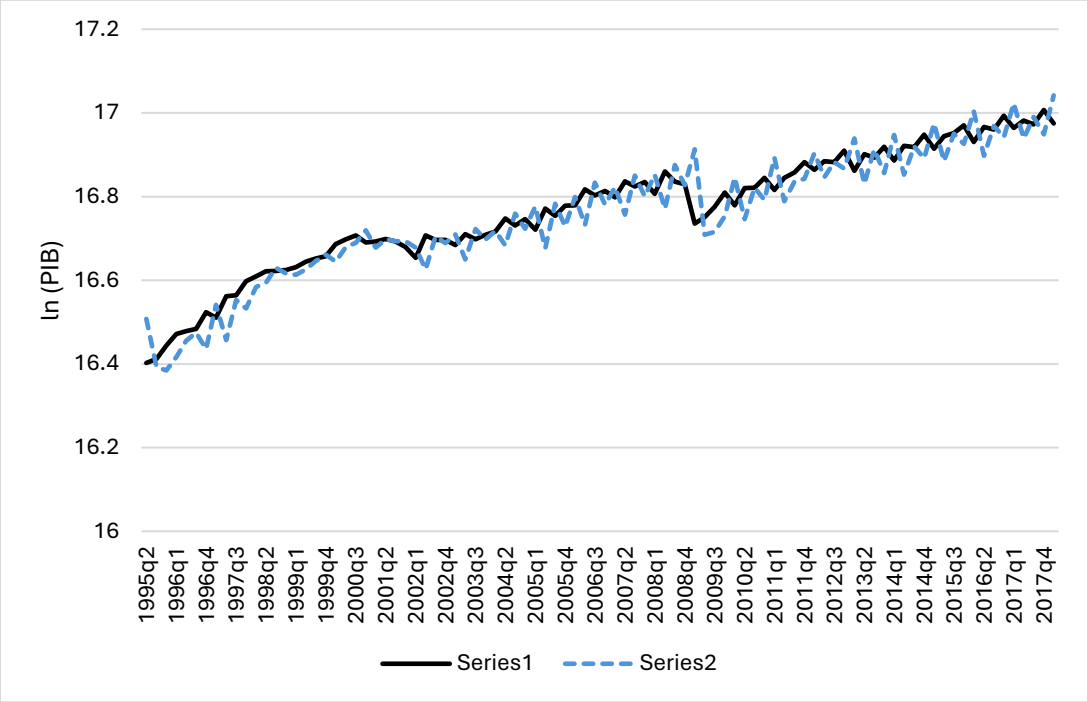
Elaboración propia

La ecuación de largo plazo muestra que el PIB está relacionado negativamente a la tasa de desempleo y positivamente al acervo de capital, lo cual coincide con la teoría económica. Por su parte la inflación es positiva al PIB, lo cual implica que del lado de la oferta el incremento en los precios se traduce en un crecimiento del PIB, para el periodo de estudio. La curva de Philips, teorizada por Samuelson y Solow, implica una relación inversa entre el desempleo y la inflación, en consecuencia, el signo esperado para la inflación es negativo para el desempleo y positivo en relación con el crecimiento económico, lo cual se muestra en la relación del vector de cointegración obtenido.

A pesar de que el coeficiente de ajuste relacionado al PIB es significativo el signo es positivo, lo cual implica que el PIB no regresa a su estado de equilibrio de largo plazo, por otro lado, el coeficiente de ajuste del Acervo de Capital es negativo y estadísticamente significativo, lo cual se traduce en un efecto estabilizador para el PIB por parte del Acervo de Capital.

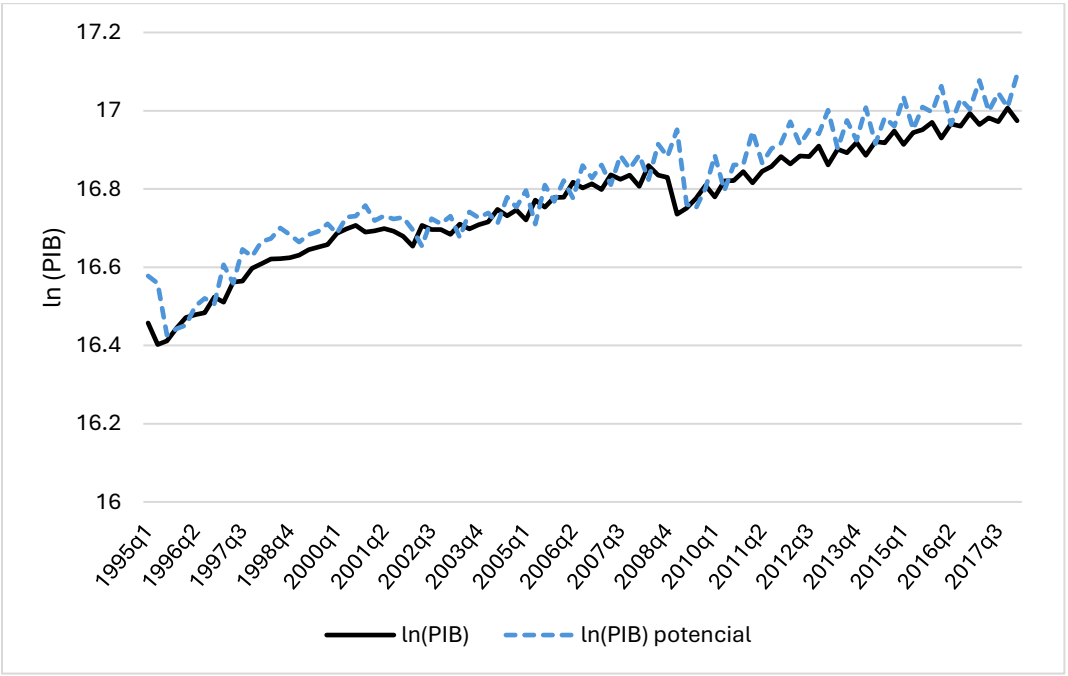
Posterior a la estimación del cambio en el PIB por el ajuste en el error se crean dos escenarios para la estimación del PIB potencial, el primero es bajo el supuesto de desempleo cero y el segundo se refiere a un incremento en el acervo de capital total en un 5%.

Gráfica 4-1: PIB estimado mediante VEC



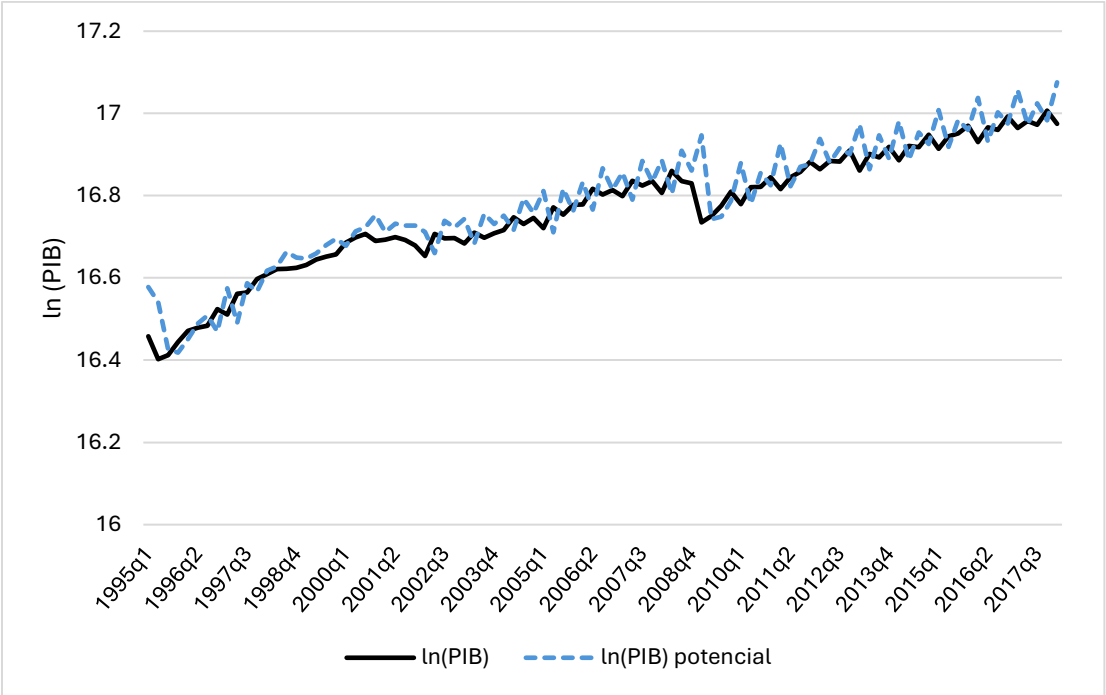
Estimación propia

Gráfica 4-2: PIB potencial estimado cuando la tasa de desempleo es 0 mediante VEC



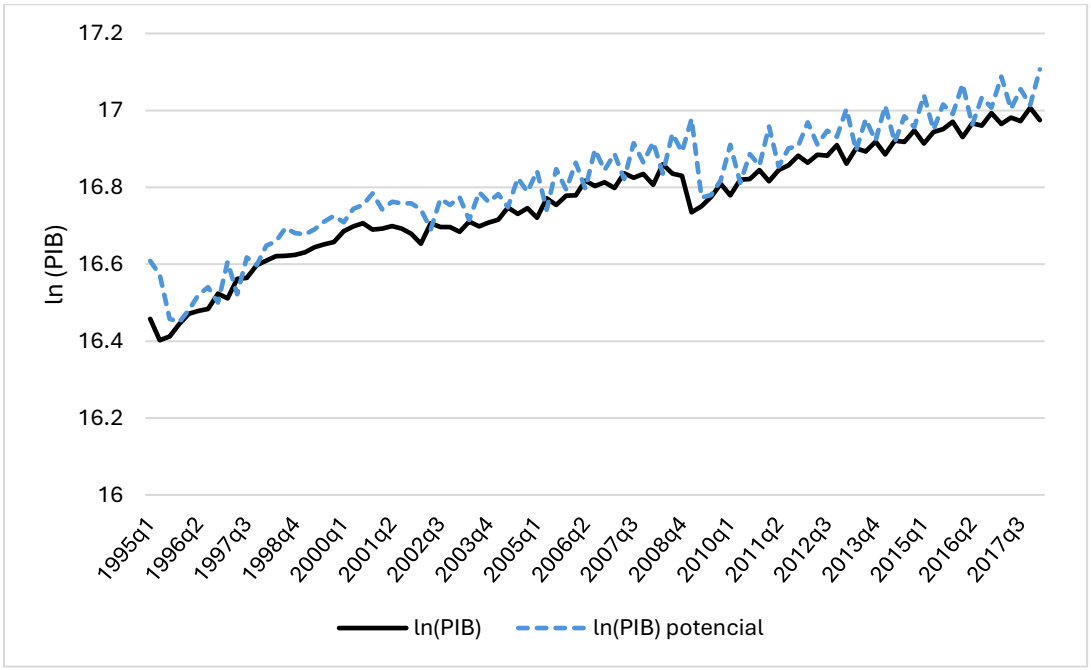
Estimación propia

Gráfica 4-3: PIB estimado cuando se incrementa el Acervo de Capital en un 10 % mediante VEC



Estimación propia

Gráfica 4-4: PIB estimado cuando se incrementa el Acervo de Capital en un 10 % mediante VEC



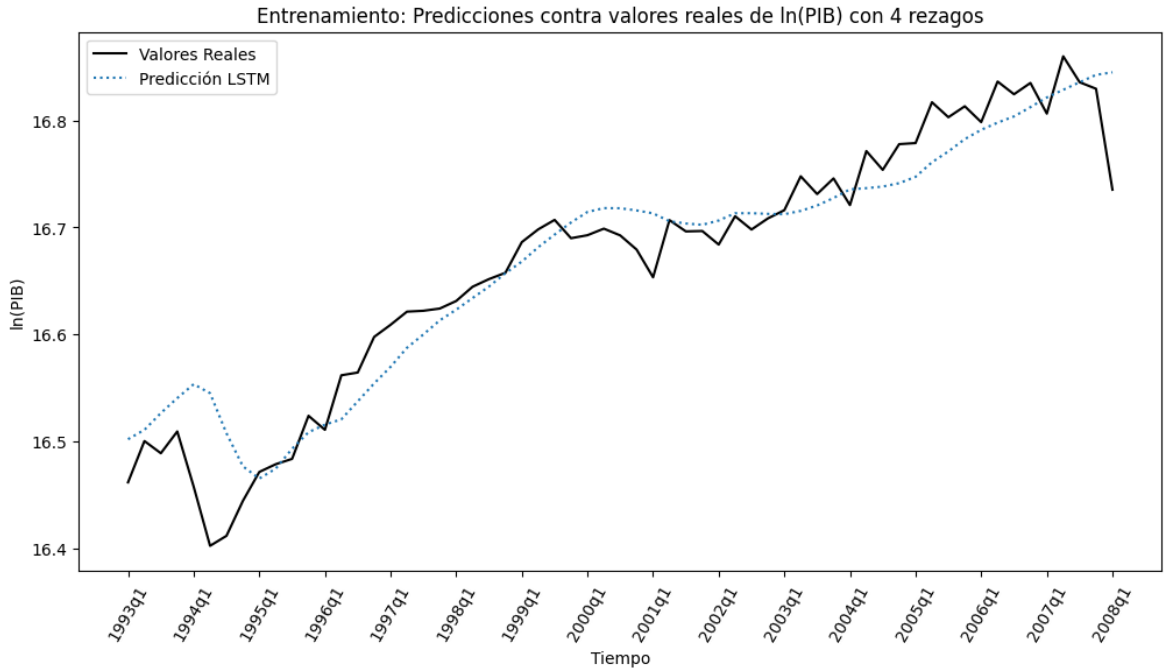
Estimación propia

Para el modelo de inteligencia artificial es necesario dividir el conjunto de información en dos grupos, uno de entrenamiento y uno de prueba, de manera que los estimadores calculados desconocen una porción de la información, permitiendo identificar si los parámetros funcionan adecuadamente con datos no conocidos. Dicha técnica permite evitar el sobreajuste, que es el equivalente a “acomodar” los parámetros para reducir la función de pérdida a un valor cercano a cero en el conjunto de datos de entrenamiento.

Se realizaron estimaciones con 1 a 6 rezagos y, dado que se selecciona el modelo con el número de rezagos que mejor se ajusta a los valores reales, se muestran los resultados de la estimación del PIB en ambos conjuntos, así mismo se muestran solamente los resultados de las estimaciones del conjunto de datos de prueba para analizar las estimaciones correspondientes al PIB potencial en los mismos escenarios que los analizados mediante el modelo VEC.

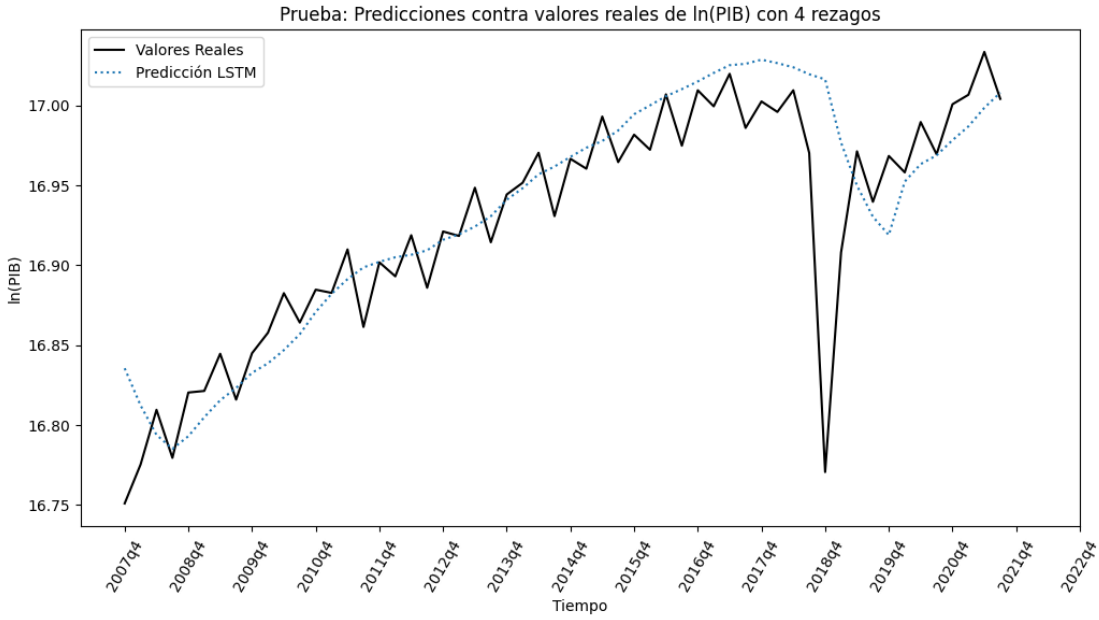
Los datos utilizados son los mismos que para el caso del modelo VEC, es decir: El PIB, la tasa de desempleo, la tasa de inflación y el Acervo de Capital total con ajuste de Almon. El detalle del código se agrega en el Apéndice D.

Gráfica 4-5: PIB estimado mediante red LSTM en conjunto de entrenamiento



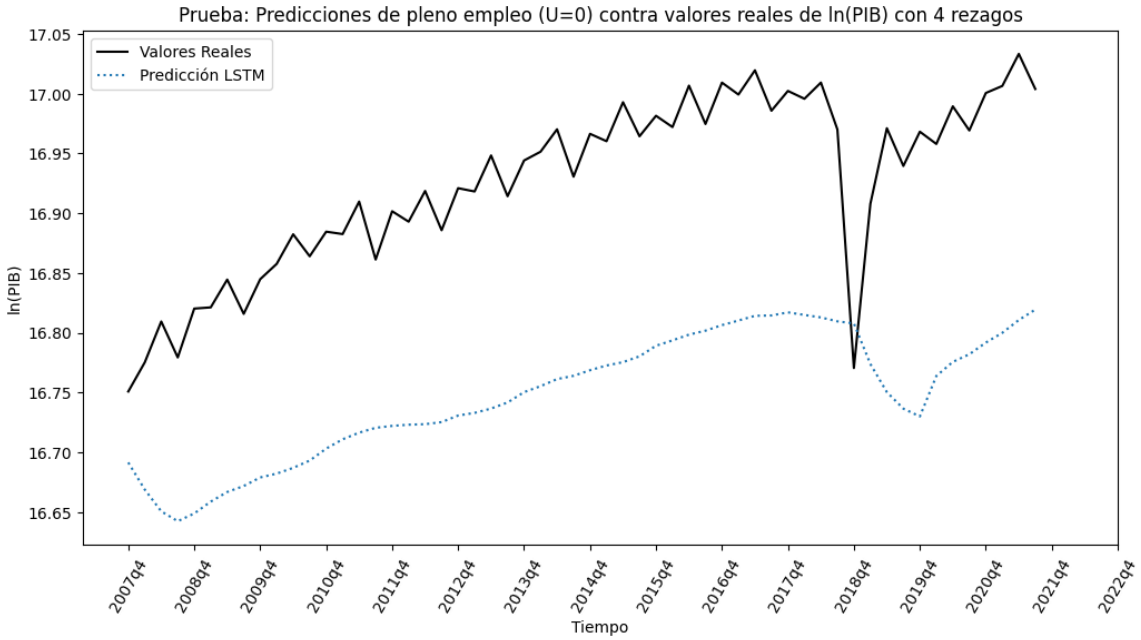
Estimación propia

Gráfica 4-6: PIB estimado mediante red LSTM en conjunto de prueba



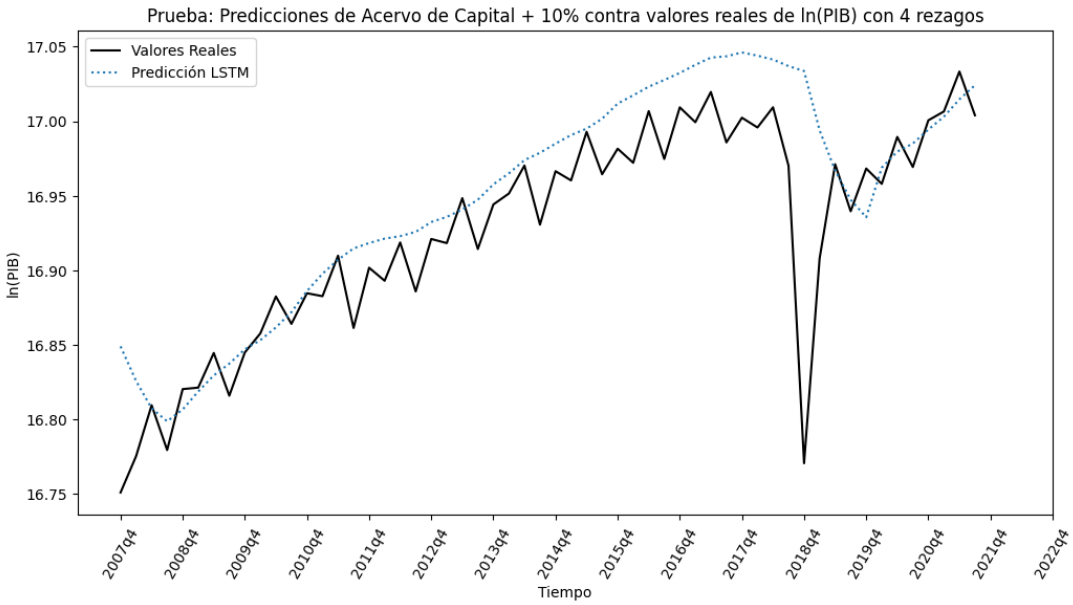
Estimación propia

Gráfica 4-7: PIB potencial estimado cuando la tasa de desempleo es cero mediante red LSTM en conjunto de prueba



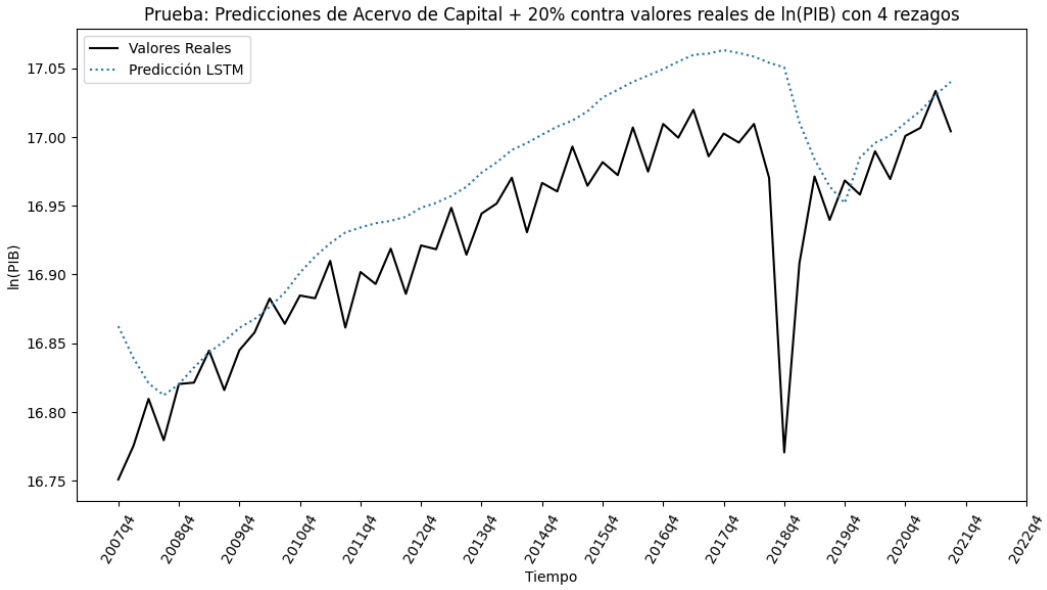
Estimación propia

Gráfica 4-8: PIB potencial estimado cuando el Acervo de Capital se incrementa en 10% red LSTM en conjunto de prueba



Estimación propia

Gráfica 4-9: PIB potencial estimado cuando el Acervo de Capital se incrementa en 20% red LSTM en conjunto de prueba



Estimación propia

La mejor estimación se logró con 4 rezagos, lo cual tiene sentido en la teoría económica ya que la red fue capaz de detectar el comportamiento estacional del PIB y no necesitó de más rezagos para mejorar su predicción. Los gráficos de las estimaciones con los demás rezagos de la red neuronal se encuentran en el Apéndice E.

Tabla 4-21: Valores de errores por número de rezago en red LSTM

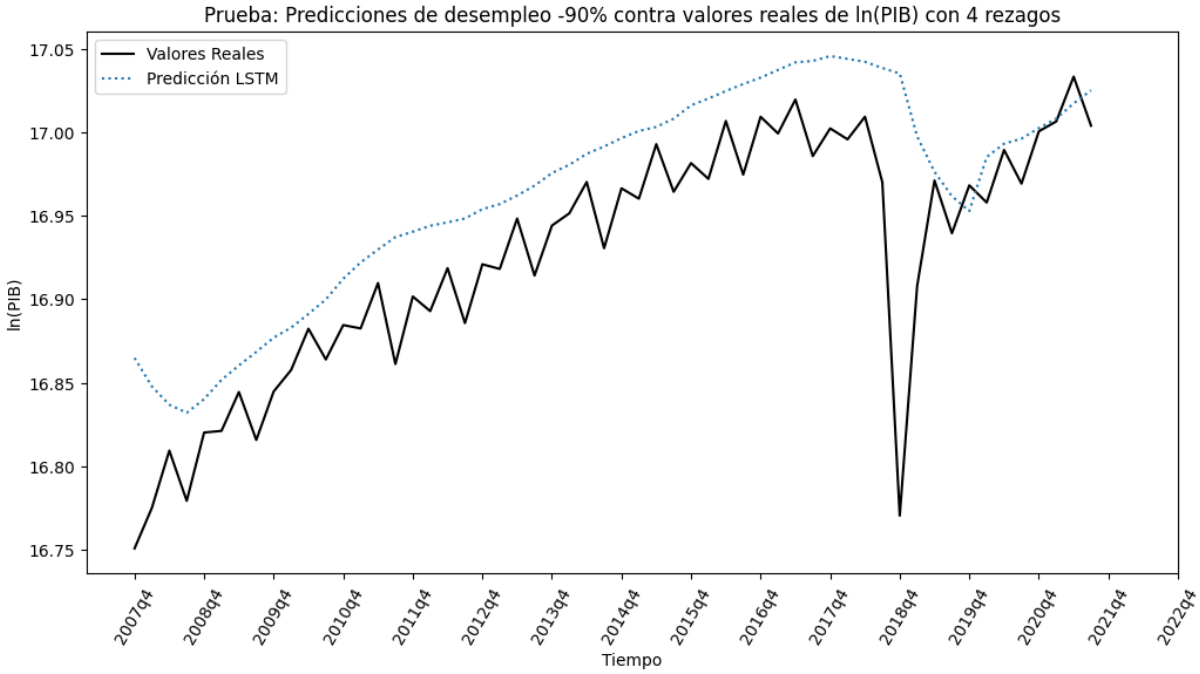
Rezagos	**PEE**	**PEP**	**ME**
1	0.03127	0.02708	0.00420
2	0.02522	0.03352	0.00830
3	0.02135	0.02649	0.00515
4	**0.02672**	**0.02395**	**0.00277**
5	0.02306	0.03659	0.01353
6	0.02313	0.03321	0.01008

Elaboración propia

Las restricciones o ajustes impuestos para el cálculo del PIB potencial (tasa de desempleo cero e incrementos en el nivel del Acervo de Capital Total), muestran efectos diferentes en cuanto a las estimaciones mediante VEC. El caso más notable es el del desempleo, el cual arroja resultados poco congruentes. Existen tres posibles explicaciones a este fenómeno, la primera es que existen variables relevantes no incluidas en el modelo, la segunda es que el modelo detecta la relación estructural con el capital y la inflación haciendo que un vector de empleo cero no le cause sentido o que considere que sería negativo para el crecimiento por los otros factores, la última es que el empleo tiene un comportamiento lineal y que se cuenta con poca información para detectar dicho comportamiento por parte de la red[8]. Con el fin de descartar un error de la red neuronal se agrega un escenario cercano al pleno empleo, es decir, una disminución del desempleo en 90%.

[8] Una red neuronal artificial requiere mucha información para detectar comportamiento lineal porque se especializa en el comportamiento no lineal.

Gráfica 4-10: PIB potencial estimado cuando el desempleo se disminuye en 90% red LSTM en conjunto de prueba



Elaboración propia

Los resultados del nuevo escenario fortalecen la explicación de la capacidad de la red para detectar un comportamiento estructural muy fuerte entre el capital y el empleo o de identificar escenarios incongruentes como desempleo cero, en ambos casos se muestra evidencia empírica de la capacidad de la red neuronal para interpretar series económicas.

En todos los casos se observa rezago en la caída de los niveles de productividad derivados de la crisis del COVID-19, lo cual se atribuye al funcionamiento de la red neuronal tipo LSTM, la cual les da diferentes pesos a periodos pasados de forma no lineal, de modo que los resultados son más fieles para proyectar tendencias. En caso de que se presente una “ruptura” o caída brusca de los valores, la estimación mantendrá por algunos periodos la inercia, posteriormente muestra la caída y corrige el comportamiento para los periodos posteriores.

Para hacer comparables los modelos, se toman las estimaciones del PIB de cada modelo y se analiza su diferencia respecto a los resultados del PIB potencial arrojado por cada escenario. Dado que los modelos estiman el logaritmo natural del PIB los porcentajes son obtenidos de la

diferencia directa de dichos logaritmos, ya que la diferencia de los logaritmos se aproxima al porcentaje en niveles de cualquier variable a la que se le aplique esta transformación.

Tabla 4-22: Porcentaje acumulado de la variación entre los resultados del PIB real y el PIB potencial para cada modelo[9]

Diferencia	**Pleno empleo**	**Capital**		**Desempleo**
		+ 10%	**+ 20%**	**-90%**
VEC	255.19%	151.30%	327.79%	244.95%
LSTM	-995.30%	132.55%	221.09%	215.59%

Elaboración propia

Sin considerar el escenario que fue incongruente para la red neuronal, los demás escenarios coinciden en la diferencia positiva entre el PIB potencial y el PIB real para cada uno de los escenarios, de hecho, la red detecta un crecimiento acumulado menor en todos los escenarios respecto a los arrojados por el modelo VEC.

Lo último podría explicarse como una mejora a los modelos econométricos tradicionales, puesto que tiene más sentido que un incremento lineal de los factores de producción no lleva necesariamente a un incremento en la misma proporción de la capacidad productiva. El resultado evidente es el incremento del Acervo de Capital, el cual incrementa en más del doble el crecimiento de la productividad (de 151.3% a 327.79%) al incrementarse el doble el porcentaje de acervo (de 10% a 20%), efecto que se observa disminuido en la red neuronal (de 132.55% a 221.09%), lo que no representa el doble del crecimiento acumulado.

La no linealidad de los factores de producción se observa en la teoría económica a través del análisis de la inversión. Un incremento excesivo del Acervo de Capital representa un contrapeso en el crecimiento económico, debido a que con una tasa de desempleo dada la subutilización del capital conlleva a una aceleración de la depreciación en niveles. La teoría de este efecto se le atribuye a Kalecki y sus análisis sobre el efecto de la inversión y la dinámica de los acervos de capital.

[9] Dado que los resultados válidos para analizar el comportamiento de la red neuronal tipo LSTM son aquellos correspondientes al conjunto de datos de prueba se homologan los periodos de tiempo para los valores presentados en la tabla.

En ese sentido podemos entender que la red neuronal muestra resultados “reales” de los escenarios de una caída en la tasa de desempleo y el incremento en el Acervo de Capital Total y que el resultado inmediato de incluir la no linealidad es ampliar el concepto estricto de máximo potencial. El modelo de redes neuronales tipo LSTM muestra evidencia de ser un potente predictor de series de tiempo económicas aún con pocos datos, respecto a la cantidad de información que suele utilizarse en la Inteligencia Artificial.

4.1. Análisis

Las estimaciones mediante métodos lineales, en específico mediante el uso de Vectores Autorregresivos y de Corrección de Errores muestran un comportamiento esperado por la teoría. Se observa una mayor sensibilidad por parte del Acervo de Capital respecto al nivel de empleo, ello puede explicarse debido a la composición del tipo de empleo en México, del cual se estima una tasa de informalidad del 55.4% y una participación en el PIB del 24.4%, Instituto Nacional de Estadística y Geografía (2023).

Por otra parte, se observa que el acervo de capital es más sensible en el modelo lineal, el cual puede explicarse cuando se toma en consideración que existe un proceso de desaceleración en el crecimiento del Acervo de Capital Total, el cual (a su vez) tiene su origen en la disminución de los flujos de Capital, es decir, de la inversión, que se observa a partir del año 2018 y tiene sus efectos en el acervo de capital a partir del año 2019 debido a la tasa de depreciación.

Uno de los aspectos más reveladores del modelo lineal es el cambio de régimen o el rompimiento estructural del vector de cointegración, según la prueba de Gregory-Hansen se da para el caso de México en el segundo cuatrimestre del 2018. Si esto se complementa con la información relacionada a la disminución de los flujos de capital el resultado es un efecto real en la capacidad productiva del país directamente relacionado a la caída en la inversión. En el escenario en el que los flujos mantienen la misma inercia, la depreciación de los bienes muebles e inmuebles se comerá la capacidad productiva del país, agudizando el problema de crecimiento, lo que en consecuencia reduce nuevamente los niveles de inversión, de este modo se cierra un círculo vicioso que se muestra como un potencial generador de una crisis sin precedentes en el país.

El modelo no lineal mediante redes neuronales muestra un comportamiento similar que es menos sensible en relación a los cambios en el desempleo y el Acervo de Capital (especialmente en este último), esto puede deberse a dos motivos probables, el primero es que los modelos de

redes neuronales requieren de mucha información para realizar estimaciones más precisas, más tratándose de series de tiempo en las que existe una única dirección y ordenamiento de los datos, a diferencia de los análisis de imágenes en las que pueden reordenarse los grupos de datos para crear conjuntos diferentes y validarse entre ellos, dejando a las series temporales con un número muy limitado de grupos de validación.

Una solución es lo que se conoce como ajuste fino (fine tuning), la cual se enfoca en encontrar los hiperparámetros que mejores resultados arrojan para un análisis particular, en el caso del PIB potencial el fine tuning se lograría mediante la aplicación del modelo propuesto en el presente estudio para otros países con características socio económicas similar a las de México, bajo el supuesto a priori de dicha similitud. Una vez aplicado el modelo a otras economías pueden extraerse los hiperparámetros que mejores resultados obtuvieron para el conjunto de los países y usarse en el modelo para el caso específico de México.

La dificultad de realizar dicho procedimiento es que deben homologarse las variables y la periodicidad de las mismas para obtener hiperparámetros válidos, si consideramos que la mayoría de los países que tienen características similares a las de México son de América Latina y que esa misma muestra tiene una serie de problemáticas que afectan directamente a la disponibilidad de la información, tal como lo es el acceso a plataformas digitales, la existencia de institutos de cuentas nacionales, las diversas prohibiciones de creación de información estadística en los periodos de dictadura, los métodos para la estimación de las variables macroeconómicas, entre otras problemáticas.

El segundo motivo probable por el que se observa una menor sensibilidad de los factores y en especial del Acervo de Capital se complementa con la teoría de Kalecki (1937) y estudios posteriores sobre la capacidad productiva ociosa o la subutilización del capital, es posible en ese sentido que la red neuronal haya sido capaz de detectar dicho comportamiento económico y por tanto haya mostrado una estimación que no es estrictamente un máximo. En otras palabras, la red neuronal está proyectando un escenario real de incrementar el Acervo de Capital y no el escenario teórico del máximo potencial debido a la correcta captura de los efectos no lineales de la subutilización del capital.

A pesar de dichas diferencias con el modelo lineal se observa claramente que un incremento en el Acervo de Capital hubiera reducido significativamente el impacto de la pandemia de COVID-19, si bien los estimados son de un nivel potencial, la evidencia del modelo no lineal muestra un

escenario más “real”, relacionado al comportamiento de la economía con los factores incrementados, siendo estos resultados un doble argumento para identificar el nivel de gravedad que representa la disminución de la inversión ante eventos de crisis económica.

Conclusiones

Los resultados obtenidos buscan fortalecer las técnicas de análisis empírico de la economía, aprovechando la capacidad actual de cálculo computacional y los modelos más recientes de análisis de series de tiempo. En dichas incorporaciones deben tomarse en cuenta las ventajas y desventajas que representa la aplicación de cada modelo.

Por ejemplo, el modelo lineal muestra una estimación del PIB potencial que resulta más congruente con la teoría económica, en especial con una función de producción que además de un sesgo (constante) tecnológico incluye al capital, el empleo y la tasa de inflación. Dicha estimación coincide con la definición de un nivel productivo máximo, el cual teóricamente puede alcanzarse durante algún periodo en el tiempo, dado el conjunto de características económicas del país a analizar.

En contraparte, el modelo no lineal mediante redes neuronales muestra escenarios que en primera instancia parecen no coincidir con la teoría pero que un análisis más profundo destaca características que deben atenderse en estudios posteriores. El modelo de redes neuronales puede mejorarse si se realiza un fine tuning de los hiperparámetros utilizando a otras economías similares a la de México, incluyendo en el conjunto de datos a las mismas variables y su correspondiente temporalidad en forma trimestral. Una vez realizado el fine tuning, del cual en el presente estudio se lleva a cabo solo de forma parcial, debe incorporarse la posibilidad de que el Acervo de Capital Total no sea relevante en la función de producción debido a una capacidad ociosa del capital en México.

Los estimadores obtenidos mediante el modelo lineal son fácilmente identificados y permiten un análisis directo sobre las características que estima el modelo de crecimiento, en contraposición, lo estimadores obtenidos por las redes neuronales son tensores que no tienen una interpretación directa y forman parte de la arquitectura del algoritmo en cuestión, en el presente estudio de las neuronas tipo LSTM. En ese sentido el modelo de redes neuronales sólo puede interpretarse de forma directa desde sus resultados, incluyendo el conjunto de métricas que establecen su validez, contrario al modelo lineal que puede interpretarse desde los estimadores propios.

En ese sentido, cabe destacar que la inteligencia Artificial ha avanzado significativamente en todos los campos y áreas del conocimiento, la economía no es la excepción. Sin duda alguna, los resultados que ofrecen los modelos de redes neuronales son una innovación en el sentido

teórico y resultados, al tiempo que permiten contrastarse con los modelos econométricos disponibles para identificar los verdaderos y escurridizos mecanismos de transmisión económicos.

Finalmente, es importante reconocer que existen ventajas y desventajas para ambos modelos, desde la perspectiva matemática relacionada a la capacidad predictiva y desde la perspectiva teórica, relacionada al análisis de las variables macroeconómicas. Lo último no limita las posibilidades de la comprensión del fenómeno económico, sino que lo complementa desde sus diversas fortalezas y áreas de oportunidad, dicho de otro modo, debe entenderse como esa sinergia entre la simplicidad que hace más eficiente el análisis teórico y la precisión de los resultados que subyacen de la complejidad de las relaciones entre variables económicas.

Bibliografía

Acevedo Rueda, R. A., Mora Mora, J. U., y Harmath Fernández, P. A. (2012). PIB potencial en Venezuela: una estimación SVAR. *Munich Personal RePEc Archive*(58691). Obtenido de https://mpra.ub.uni-muenchen.de/58691/

Agrawal, A., Gans, J., y Goldfarb, A. (2019). Introduction to "The Economics of Artificial Intelligence: An. En A. Agrawal, J. Gans, y A. Goldfarb, *The Economics of Artificial Intelligence: An Agenda* (págs. 1-19). University of Chicago Press. Obtenido de http://www.nber.org/chapters/c14005

Albrieu, R., Rapetti, M., Brest López, C., Larroulet, P., y Sorrentino, A. (octubre de 2018). Inteligencia artificial y crecimiento económico. Oportunidades y desafíos para México. *Inteligencia Artificial y Crecimiento Económico en América Latina*.

Anaya Narváez, A. R., Torres Hoyos, F., y Pinedo López, J. (diciembre de 2016). Crecimiento económico: Un modelo de vectores autorregresivos para el caso colombiano. *Revista de Ciencias Sociales, 22*(4), 41-55.

Bierens, H. J. (diciembre de 2010). Information Criteria and Model Selection. *Research Gate*. Obtenido de https://www.researchgate.net/publication/267374496_Information_Criteria_and_Model_Selection

Bjørnland, H. C. (octubre de 2000). *VAR Models in Macroeconomic.* Recuperado el 2023 de mayo de 23, de Statistics Norway: https://www.ssb.no/a/histstat/doc/doc_200014.pdf

Bokhari, S., y Geltner, D. (mayo de 2016). Characteristics of Depreciation in Commercial. *Real Estate Economics, 46*(4), 745-782. doi: http://dx.doi.org/10.1111/1540-6229.12156

Bouza Herrera, C. N. (2018). *Modelos de Regresión y sus Aplicaciones.* (GRASHUMEDIA, Ed.) Obtenido de Research Gate: https://www.researchgate.net/profile/Carlos-Bouza/publication/323227561_MODELOS_DE_REGRESION_Y_SUS_APLICACIONES/links/5a871265a6fdcc6b1a3abe40/MODELOS-DE-REGRESION-Y-SUS-APLICACIONES.pdf

Chen, K., y Papanyan, S. (2016). *In Search of Potential GDP.* BBVA, U.S. Economic Watch. Recuperado el Noviembre de 2023, de https://www.bbvaresearch.com/wp-content/uploads/2016/06/160608_US_Potential_GDP.pdf

Clark, M. A., Choi, J., y Douglas, M. (2020). *Biology* (2da ed.). OpenStax.

Clark, P. B., y Laxton, D. (febrero de 1997). Phillips Curves, Phillips Lines and the Unemployment Costs of Overheating. *Working Paper of the International Monetary Fund, 97*(17).

Corona, F., Orraca, P., y López Pérez, J. (diciembre de 2022). Evaluación de diferentes métodos para la estimación del PIB potencial el caso de México. *Estudios económicos, 37*(2), 285-313. doi:https://doi.org/10.24201/ee.v37i2.432

Damasevicius, R. (2023). *Economic Analysis Letters, 2*(2), 52-59.

Domar, E. D. (Marzo de 1947). Expansion and Employment. *The American Economic Review, 37*(1), 34-55.

Ernst, E., Merola, R., y Samaan, D. (2018). The economics of artificial intelligence: Implications for the future of work. (I. L. Office, Ed.) *ILO Future of work research paper series, 5*.

Erumban, A. A. (june de 2008). Lifetimes of machinery and equipment: evidence from dutch manufacturing. *Review of Income and Wealth, 54*(2).

Evans, G. R. (1997). *CHAPTER 1 ECONOMIC MODELS.* Recuperado el 10 de May de 2023, de HMC Courses offered by professor Gary Evans: https://www.palmislandtraders.com/econ53/chap1.pdf

Faria, M., y Leibovici, F. (Julio de 2023). Artificial Intelligence and Inflation Forecasts. *Economic Research*. doi:https://doi.org/10.20955/wp.2023.015

Felipe, J., y McCombie, J. (2007). Is a theory of total factor productivity really needed. *Metroconomica, 58*(1), 195-229.

Fisher, J. D., Stern, J., Smith, B. C., y Webb, B. (2005). Analysis of Economic Depreciation for Multi-Family Property. *Journal of Real Estate Research, 27*(4), 355-370.

Fredric, H. M., y Kostanic, I. (2001). *Principles of neurocomputing for science and engineering.* New York: McGraw Hill.

Gallardo Del Angel, R. (Marzo de 2020). Financial time series forecasting using Artificial Neural Networks. *Revista Mexicana de Economía y Finanzas, 15*(1), 105-122. doi:https://doi.org/10.21919/remef.v15i1.376

Gers, F., Cummins, F., Fernandez, S., Bayer, J., Wierstra, D., Togelius, J., . . . Graves, A. (27 de agosto de 2015). *Understanding LSTM Networks*. Obtenido de colah's blog: https://colah.github.io/posts/2015-08-Understanding-LSTMs/

Gonzalez, S. (Julio de 2000). *Neural Networks for Macroeconomic Forecasting: A Complementary Approach to Linear Regression Models.* Recuperado el Noviembre de 2023, de Gouvernement du Canada: https://publications.gc.ca/Collection/F21-8-2000-7E.pdf

Gonzalo, J., y Granger, C. (Enero de 1995). Estimation of Common Long-Memory Components in Cointegrated Systems. *Journal of Business y Economic Statistics, 13*(1), 27-35.

Goulet Coulombe, P. (Octubre de 2022). Forecasting and Understanding US Inflation with Artificial Intelligence. *SUERF Policy Brief*(455).

Greene, W. H. (2002). Capítulo 19: Modelos con variables rezagadas. En W. H. Greene, *Análisis Econométrico.* Nueva Jersey: Prentice Hall.

Gregory, A. W., y Hansen, B. E. (1996). Residual-bsaed tests for cointegration in models with regime shifts. *Journal of econometrics*, 99-126.

Gutierrez Cruz, F. S., y Moreno Brid, J. C. (diciembre de 2022). Estimación del Acervo de Capital público y privado en México a nivel estatal (2004-2018). *Análisis Económico, 37*(96), 161-180.

Harrod, H. R. (Marzo de 1939). An Essay in Dynamic Theory. *The Economic Journal, 49*(193), 14-33.

Herbrich, R., Keilbach, M., Graepel, T., Bollmann-Sdorra, P., y Obermayer, K. (2012). Neural Networks in Economics. En *Computational Techniques for Modelling Learning in Economics* (págs. 169-196). Springer US.

Huang, W., Keung Lai, K., Nakamori, Y., Wang, S., y Yu, L. (2007). Neural Networks in Finance and Economics Forecasting. *International Journal of Information Technology y Decision Making, 6*(1), 113-140. doi:10.1142/S021962200700237X

Instituto Nacional de Estadística y Geografía. (2023). Medición de la economía informal. *Comunicado de prensa número 824/23*, págs. 1-10. Instituto Nacional de Estadística y Geografía. Obtenido de https://www.inegi.org.mx/contenidos/saladeprensa/boletines/2023/MDEI/MDEI2022.pdf

International Telecommunication Union. (2018). *Assessing the Economic Impact of Artificial Intelligence.* Geneva: International Telecommunication Union. Obtenido de https://www.itu.int/dms_pub/itu-s/opb/gen/S-GEN-ISSUEPAPER-2018-1-PDF-E.pdf

Jahan, S., y Mahmud, A. S. (Junio de 2015). ¿Qué es el capitalismo? *Finanzas y Desarrollo, 52*(2), 44-45.

Johansen, S. (octubre de 1987). Statistical Analysis of Cointegration Vectors. *Institute of Mathematical Statistics University of Copenhagen, Preprint 1987*(7).

Jurafsky, D., y Martin, J. H. (Febrero de 2024). RNNs and LSTMs. En D. Jurafsky, y J. H. Martin, *Speech and Language Processing* (3rd ed.). Obtenido de https://web.stanford.edu/~jurafsky/slp3/

Kaastra, I., y Boyd, M. (Marzo de 1995). Designing a neural network for forecasting financial and economic time series. *Elsevier, Neurocomputing*(10), 215-236.

Kalecki, M. (febrero de 1937). A Theory of the Business Cycle. *The Review of Economic Studies, 4*(2), 77-97.

Karanassou, M., Sala, H., y Snower, D. J. (agosto de 2006). Phillips Curves and Unemployment Dynamics: A Critique and a Holistic Perspective. *IZA discussion paper*(2265).

Keynes, J. M. (1936). *General theory of employment interest and money* (2008 ed.). ETH Zurich.

Klein, D. B., y Tullok, G. (2001). *A Plea to Economics Who Favour Liberty: Assist the Everyman (Occasional Paper, 118).* Institute of Economic Affairs.

Levy, D. M., y Feigenbaum, S. (Septiembre de 1990). Testing the replication hypothesis When the Data Set is Subject to Gross Error. *Economics Letters, 34*(1), 49-53.

Liquet, B., Moka, S., y Nazarathy, Y. (2024). *Mathematical Engineering of Deep Learning.* CRC Press.

Lütkepohl, H. (2005). *Nueva introducción al análisis de series temporales múltiples.* Alemania: Springer.

Lütkepohl, H. (2007). *Análisis ecnométrico con modelos de vectores autorregresivos.* Italia: Departamento de economía del Instituto de la Universidad Europea.

Malthus, T. (1798). *An Essay on the Principle of Population* (1998 ed.). The Electronic Scholarly Publishing Project.

Marx, K. (1867). *El capital tomo I* (2008 ed.). Siglo veintiuno.

McCloskey, D. N. (Mayo de 1985). The Loss Function has been mislaid: the Rhetoric of Significance Tests. *American Economic Review, 75*(2), 201-205.

Mehtab, S., Sen, H., y Dutta, A. (septiembre de 2020). Stock Price Prediction Using Machine Learning and LSTM-Based Deep Learning Models. *arXiv*. doi:https://doi.org/10.1007/978-981-16-0419-5_8

Panchal, F. S., y Panchal, M. (Noviembre de 2014). Review on Methods of Selecting Number of Hidden Nodes in Artificial Neural Network. *International Journal of Computer Science and Mobile Computing, 3*(11), 455-464.

Pilström, P., y Pohl, S. (2009). *Forecasting GDP Growth: The case of The Baltic States.* Sweden: Jönköping International Business School.

Roush, J., Siopes, K., y Hu, G. (junio de 2017). Predicting gross domestic product using autoregressive models. *2017 IEEE 15th International Conference on Software Engineering Research, Management and Applications (SERA).* doi:10.1109/SERA.2017.7965745

Say, J. (1880). *A Treatise on Political Economy* (2001 ed.). C.R. Prinsep, M.A.

Sazli, M. H. (enero de 2006). A brief review of feed-forward neural networks. *Communications Faculty of Sciences University of Ankara Series A2-A3 Physical Sciences and Engineering, 50*(1), 11-17. doi: 10.1501/commua1-2_0000000026

Smith, A. (1784). *An inquiry into the nature and causes of the wealth of nations* (2007 ed.). MetaLibri.

Solow, R. M. (Febrero de 1956). A Contribution to the Theory of Economic Growth. *The Quarterly Journal of Economics, 70*(1), 65-94.

Sossa Azuela, J. H. (2020). El papel de la inteligencia artificial en la Industria 4.0. En P. O. Rodríguez Reséndiz, *Inteligencia artificial y datos masivos en archivos digitales sonoros y audiovisuales* (págs. 21-58). UNAM.

Staudemeyer, R. C., y Rothstein Morris, E. (septiembre de 2019). Understanding LSTM: a tutorial into Long Short-Term Memory Recurrent Neural Networks. *arXiv*. doi:https://doi.org/10.48550/arXiv.1909.09586

Sun, R. (diciembre de 2019). Optimization for deep learning: theory and algorithms. *arXiv*. doi:https://doi.org/10.48550/arXiv.1912.08957

Szczepański, M. (Julio de 2019). Economic impacts of artificial intelligence (AI). *European Parliamentary Research Service*.

The Congress of the United States . (2004). *A Summary of Alternative Methods for Estimating Potential GDP.* Congressional Budget Office. Recuperado el Noviembre de 2023, de https://www.cbo.gov/sites/default/files/108th-congress-2003-2004/reports/03-16-gdp.pdf

Valle-Cruz, D., Fernandes-Cortez, V., y Gil-García, J. R. (2023). Inteligencia artificial y planeación presupuestaria en México: promesas y retos en América Latina para la asignación del gasto público. *Revista del CLAD Reforma y Democracia*(85), 5-52.

Villada, F., Muñoz, N., y García, E. (2012). Aplicación de las Redes Neuronales al Pronóstico de Precios en el Mercado de Valores. *Información Tecnológica, 23*(4), 11-20. doi:http://dx.doi.org/10.4067/S0718-07642012000400003

Wooldridge, J. M. (2013). Regression Analysis with Cross-Sectional Data. En J. M. Wooldridge, *Introductory econometrics: A modern approach* (5th ed.). South-Western.

Wulwick, N. (1992). Kaldor's Growth Theory. *Journal of the History of Economic Thought, 14*(1), 36-54.

Apéndice

Apéndice A

Estimaciones de autorregresión vectorial

Muestra (ajustada): 1995T2 2018T1

Errores estándar en ( ) y estadísticos t en [ ]

	LPIB	**U**	**LAKTA**	**INF_AN**
LPIB(-1)	0.389130	-0.052499	-0.01222	0.132052
	(0.12991)	(0.03001)	(0.02639)	(0.11885)
	[2.99528]	[-1.74938]	[-0.46311]	[1.11108]
LPIB(-2)	0.293359	0.013037	0.024470	-0.054507
	(0.14311)	(0.03306)	(0.02907)	(0.13092)
	[2.04996]	[0.39438]	[0.84186]	[-0.41635]
LPIB(-3)	-0.024426	0.013183	-0.037509	0.113233
	(0.12821)	(0.02962)	(0.02604)	(0.11729)
	[-0.19052]	[0.44512]	[-1.44040]	[0.96542]
LPIB(-4)	0.621847	-0.023142	0.021760	-0.418846
	(0.12810)	(0.02959)	(0.02602)	(0.11719)
	[4.85430]	[-0.78204]	[0.83631]	[-3.57402]
LPIB(-5)	-0.174731	0.067825	0.071964	0.078368
	(0.14712)	(0.03399)	(0.02988)	(0.13459)
	[-1.18765]	[1.99572]	[2.40821]	[0.58226]
LPIB(-6)	-0.231086	-0.060081	-0.029846	0.025298
	(0.14499)	(0.03349)	(0.02945)	(0.13264)
	[-1.59383]	[-1.79389]	[-1.01348]	[0.19073]

	LPIB	U	LAKTA	INF_AN
LPIB(-7)	0.277568	0.007733	0.026935	0.069508
	(0.13431)	(0.03103)	(0.02728)	(0.12287)
	[2.06658]	[0.24923]	[0.98734]	[0.56568]
LPIB(-8)	0.115691	0.015877	-0.001643	0.143648
	(0.14165)	(0.03272)	(0.02877)	(0.12959)
	[0.81671]	[0.48520]	[-0.05710]	[1.10848]
LPIB(-9)	-0.086552	-0.01169	0.007273	-0.141776
	(0.11275)	(0.02605)	(0.02290)	(0.10315)
	[-0.76763]	[-0.44884]	[0.31757]	[-1.37448]
U(-1)	0.761557	0.580521	-0.021344	-0.301016
	(0.65755)	(0.15189)	(0.13356)	(0.60155)
	[1.15817]	[3.82188]	[-0.15981]	[-0.50040]
U(-2)	-0.277528	0.205328	-0.170567	0.601925
	(0.77153)	(0.17822)	(0.15671)	(0.70582)
	[-0.35971]	[1.15208]	[-1.08844]	[0.85280]
U(-3)	-0.16291	0.014307	0.100328	-1.220961
	(0.70667)	(0.16324)	(0.14353)	(0.64648)
	[-0.23053]	[0.08764]	[0.69899]	[-1.88862]
U(-4)	-0.324359	0.349049	0.199316	0.093270
	(0.69049)	(0.15950)	(0.14025)	(0.63169)
	[-0.46975]	[2.18835]	[1.42117]	[0.14765]

	LPIB	U	LAKTA	INF_AN
U(-5)	0.506497	-0.119258	-0.011422	1.109564
	(0.71480)	(0.16512)	(0.14518)	(0.65392)
	[0.70859]	[-0.72226]	[-0.07867]	[1.69679]
U(-6)	0.224700	0.007362	-0.060906	-0.023966
	(0.74918)	(0.17306)	(0.15217)	(0.68537)
	[0.29993]	[0.04254]	[-0.40026]	[-0.03497]
U(-7)	-0.752796	0.045009	0.113705	1.033159
	(0.70597)	(0.16308)	(0.14339)	(0.64584)
	[-1.06633]	[0.27600]	[0.79297]	[1.59971]
U(-8)	0.360135	0.097822	0.122908	-1.707838
	(0.70505)	(0.16287)	(0.14320)	(0.64500)
	[0.51080]	[0.60063]	[0.85827]	[-2.64782]
U(-9)	0.861317	-0.296831	0.100743	0.459085
	(0.56390)	(0.13026)	(0.11454)	(0.51588)
	[1.52742]	[-2.27873]	[0.87957]	[0.88991]
LAKTA(-1)	1.226928	0.010497	1.526595	-2.088865
	(0.79220)	(0.18300)	(0.16091)	(0.72473)
	[1.54876]	[0.05736]	[9.48749]	[-2.88226]
LAKTA(-2)	-2.794814	0.047306	-0.626411	3.791584
	(1.21372)	(0.28037)	(0.24652)	(1.11035)
	[-2.30268]	[0.16873]	[-2.54099]	[3.41476]

	LPIB	U	LAKTA	INF_AN
LAKTA(-3)	1.255856	-0.040024	-0.163822	-3.452622
	(1.10661)	(0.25563)	(0.22477)	(1.01236)
	[1.13486]	[-0.15657]	[-0.72885]	[-3.41046]
LAKTA(-4)	-0.32363	0.261919	0.647427	4.108481
	(1.11933)	(0.25856)	(0.22735)	(1.02399)
	[-0.28913]	[1.01298]	[2.84773]	[4.01222]
LAKTA(-5)	-1.415813	-0.405267	-1.045421	-1.358544
	(1.17724)	(0.27194)	(0.23911)	(1.07697)
	[-1.20266]	[-1.49028]	[-4.37212]	[-1.26145]
LAKTA(-6)	4.151945	0.213797	0.636436	-2.808816
	(1.15072)	(0.26581)	(0.23372)	(1.05271)
	[3.60814]	[0.80431]	[2.72302]	[-2.66818]
LAKTA(-7)	-3.354809	0.009342	-0.108756	3.715854
	(1.12218)	(0.25922)	(0.22793)	(1.02660)
	[-2.98956]	[0.03604]	[-0.47715]	[3.61957]
LAKTA(-8)	1.412375	-0.210257	0.273057	-3.451515
	(1.15133)	(0.26596)	(0.23385)	(1.05327)
	[1.22674]	[-0.79057]	[1.16766]	[-3.27696]
LAKTA(-9)	-0.431926	0.141340	-0.242734	1.569086
	(0.50771)	(0.11728)	(0.10312)	(0.46447)
	[-0.85073]	[1.20513]	[-2.35384]	[3.37822]

	LPIB	U	LAKTA	INF_AN
INF_AN(-1)	-0.112003	0.043191	0.005679	1.143118
	(0.11763)	(0.02717)	(0.02389)	(0.10761)
	[-0.95220]	[1.58956]	[0.23770]	[10.6230]
INF_AN(-2)	-0.062733	-0.074139	-0.057271	-0.232132
	(0.18127)	(0.04187)	(0.03682)	(0.16583)
	[-0.34608]	[-1.77058]	[-1.55552]	[-1.39981]
INF_AN(-3)	0.167043	0.037795	0.078867	0.238077
	(0.16996)	(0.03926)	(0.03452)	(0.15548)
	[0.98284]	[0.96267]	[2.28462]	[1.53121]
INF_AN(-4)	-0.19903	0.002370	-0.073792	-0.384285
	(0.17044)	(0.03937)	(0.03462)	(0.15592)
	[-1.16777]	[0.06020]	[-2.13163]	[-2.46463]
INF_AN(-5)	-0.033051	0.001790	-0.015185	0.215275
	(0.15565)	(0.03595)	(0.03161)	(0.14239)
	[-0.21235]	[0.04977]	[-0.48032]	[1.51188]
INF_AN(-6)	0.253114	-0.050225	0.007339	-0.054017
	(0.14775)	(0.03413)	(0.03001)	(0.13517)
	[1.71310]	[-1.47156]	[0.24455]	[-0.39963]
INF_AN(-7)	-0.118057	0.024633	0.048007	0.222446
	(0.15373)	(0.03551)	(0.03122)	(0.14064)
	[-0.76795]	[0.69366]	[1.53749]	[1.58171]

	LPIB	U	LAKTA	INF_AN
INF_AN(-8)	0.167199	-0.009177	-0.021513	-0.528881
	(0.11332)	(0.02618)	(0.02302)	(0.10367)
	[1.47548]	[-0.35057]	[-0.93467]	[-5.10172]
INF_AN(-9)	-0.184423	0.013588	-0.02445	0.258166
	(0.07018)	(0.01621)	(0.01425)	(0.06420)
	[-2.62780]	[0.83812]	[-1.71520]	[4.02101]
C	1.917703	-0.016022	0.678453	0.448209
	(1.14285)	(0.26400)	(0.23213)	(1.04551)
	[1.67801]	[-0.06069]	[2.92278]	[0.42870]
D09Q1	-0.091985	0.006201	-0.004749	-0.002125
	(0.01292)	(0.00298)	(0.00262)	(0.01182)
	[-7.12141]	[2.07840]	[-1.81004]	[-0.17984]
D98Q4	-0.013661	0.003075	0.000298	0.054013
	(0.01369)	(0.00316)	(0.00278)	(0.01253)
	[-0.99752]	[0.97203]	[0.10706]	[4.31121]
D05Q1	-0.028384	0.008538	-0.001278	-0.000568
	(0.01148)	(0.00265)	(0.00233)	(0.01050)
	[-2.47337]	[3.22077]	[-0.54833]	[-0.05411]
D98Q1	0.030093	0.006331	0.012556	0.025450
	(0.02058)	(0.00475)	(0.00418)	(0.01883)
	[1.46207]	[1.33151]	[3.00346]	[1.35162]

	LPIB	U	LAKTA	INF_AN
D99Q4	0.021011	0.000833	0.005547	-0.036739
	(0.01394)	(0.00322)	(0.00283)	(0.01276)
	[1.50687]	[0.25875]	[1.95877]	[-2.88017]
D09Q2	-0.076863	0.002460	-0.000541	-0.00604
	(0.01465)	(0.00338)	(0.00298)	(0.01340)
	[-5.24731]	[0.72713]	[-0.18180]	[-0.45075]
D02Q1	-0.057318	0.004554	-0.003811	0.008885
	(0.01230)	(0.00284)	(0.00250)	(0.01125)
	[-4.66012]	[1.60285]	[-1.52531]	[0.78966]
D08Q1	-0.035532	0.001707	-0.001051	0.008676
	(0.01221)	(0.00282)	(0.00248)	(0.01117)
	[-2.90997]	[0.60511]	[-0.42377]	[0.77669]

R2 ajustada	0.9953	0.9703	0.9998	0.9915
Suma de residuales	0.0046	0.0002	0.0002	0.0039
Estadístico- F	436.5744	68.5773	10523.3100	243.4673
Log-Verosimilitud	324.5367	459.3482	471.1843	332.7262
Akaike (AIC)	-6.0769	-9.0076	-9.2649	-6.2549
Schwarz (SC)	-4.8434	-7.7741	-8.0314	-5.0214

Apéndice B

Estimaciones de corrección de errores vectoriales

Muestra (ajustada): 1995T2 2018T1

Observaciones incluidas: 92 después de los ajustes

Errores estándar en ( ) y estadísticos t en [ ]

Intervalo de retardos (en primeras diferencias): 1 a 8

Variables endógenas: LPIB U LAKTA INF_AN

Variables exógenas (solo a corto plazo): DD09Q1 DD98Q4 DD05Q1 DD98Q1 DD99Q4 DD09Q2

Supuestos deterministas: Caso 3: La dinámica a corto plazo incluye una constante.

Cointegrating Eq:	**CointEq1**
LPIB(-1)	1.000000
U(-1)	5.066768
	(0.74530)
	[6.79827]
LAKTA(-1)	-1.402764
	(0.11768)
	[-11.9206]
INF_AN(-1)	-0.597522
	(0.21011)
	[-2.84387]

Corrección de errores:	D(LPIB)	D(U)	D(LAKTA)	D(INF_AN)
COINTEQ1	0.253680	-0.032794	0.085051	-0.06152
	(0.12033)	(0.02230)	(0.01841)	(0.10195)
	[2.10825]	[-1.47036]	[4.61864]	[-0.60342]
D(LPIB(-1))	-0.306564	-0.032222	-0.069295	0.217356
	(0.17988)	(0.03334)	(0.02753)	(0.15241)
	[-1.70424]	[-0.96640]	[-2.51718]	[1.42610]
D(LPIB(-2))	-0.095297	-0.00857	-0.033966	0.024833
	(0.18956)	(0.03514)	(0.02901)	(0.16061)
	[-0.50272]	[-0.24391]	[-1.17082]	[0.15461]
D(LPIB(-3))	-0.497957	0.026946	-0.098828	0.269324
	(0.17053)	(0.03161)	(0.02610)	(0.14448)
	[-2.92013]	[0.85251]	[-3.78698]	[1.86403]
D(LPIB(-4))	0.237282	-0.006323	-0.072371	-0.148178
	(0.19370)	(0.03590)	(0.02964)	(0.16412)
	[1.22503]	[-0.17612]	[-2.44144]	[-0.90288]
D(LPIB(-5))	-0.265602	0.077649	-0.030589	-0.009829
	(0.21016)	(0.03895)	(0.03216)	(0.17807)
	[-1.26379]	[1.99331]	[-0.95107]	[-0.05520]
D(LPIB(-6))	-0.375086	-0.010927	-0.068634	0.062461
	(0.17228)	(0.03193)	(0.02636)	(0.14597)
	[-2.17722]	[-0.34218]	[-2.60321]	[0.42790]

Corrección de errores:	**D(LPIB)**	**D(U)**	**D(LAKTA)**	**D(INF_AN)**
D(LPIB(-7))	0.038367	0.014789	-0.022537	0.067242
	(0.16176)	(0.02998)	(0.02476)	(0.13706)
	[0.23719]	[0.49326]	[-0.91041]	[0.49062]
D(LPIB(-8))	0.027500	0.024368	-0.020898	0.072681
	(0.13875)	(0.02572)	(0.02123)	(0.11756)
	[0.19820]	[0.94754]	[-0.98418]	[0.61826]
D(U(-1))	-1.615954	0.049765	-0.450369	0.534433
	(1.01506)	(0.18815)	(0.15534)	(0.86005)
	[-1.59198]	[0.26450]	[-2.89921]	[0.62140]
D(U(-2))	-1.149532	0.094908	-0.581706	0.766641
	(1.07426)	(0.19912)	(0.16440)	(0.91021)
	[-1.07007]	[0.47664]	[-3.53832]	[0.84227]
D(U(-3))	-1.35112	0.093385	-0.495179	0.085667
	(1.07433)	(0.19913)	(0.16441)	(0.91027)
	[-1.25764]	[0.46896]	[-3.01181]	[0.09411]
D(U(-4))	-2.222189	0.494731	-0.385363	0.517555
	(1.09274)	(0.20254)	(0.16723)	(0.92587)
	[-2.03360]	[2.44258]	[-2.30439]	[0.55899]
D(U(-5))	-0.746256	0.307111	-0.396015	1.471563
	(0.98214)	(0.18204)	(0.15030)	(0.83216)
	[-0.75983]	[1.68701]	[-2.63475]	[1.76837]

Corrección de errores:	**D(LPIB)**	**D(U)**	**D(LAKTA)**	**D(INF_AN)**
D(U(-6))	-1.78426	0.374397	-0.495362	0.885577
	(0.95011)	(0.17611)	(0.14540)	(0.80502)
	[-1.87795]	[2.12596]	[-3.40682]	[1.10007]
D(U(-7))	-0.721924	0.215602	-0.253873	1.315405
	(0.87535)	(0.16225)	(0.13396)	(0.74168)
	[-0.82472]	[1.32882]	[-1.89511]	[1.77355]
D(U(-8))	-0.207807	0.347702	-0.084344	-0.716295
	(0.64906)	(0.12031)	(0.09933)	(0.54994)
	[-0.32017]	[2.89015]	[-0.84912]	[-1.30249]
D(LAKTA(-1))	0.575633	0.006656	0.659445	-1.82794
	(0.90850)	(0.16839)	(0.13903)	(0.76976)
	[0.63361]	[0.03952]	[4.74301]	[-2.37467]
D(LAKTA(-2))	-2.21802	0.135778	-0.187036	2.525597
	(0.83393)	(0.15457)	(0.12762)	(0.70658)
	[-2.65973]	[0.87841]	[-1.46555]	[3.57441]
D(LAKTA(-3))	1.442672	-0.13541	-0.013329	-1.992463
	(0.89303)	(0.16553)	(0.13667)	(0.75666)
	[1.61547]	[-0.81805]	[-0.09753]	[-2.63324]
D(LAKTA(-4))	-1.031516	0.316449	0.467458	2.412479
	(0.81653)	(0.15135)	(0.12496)	(0.69184)
	[-1.26329]	[2.09088]	[3.74087]	[3.48706]

Corrección de errores:	**D(LPIB)**	**D(U)**	**D(LAKTA)**	**D(INF_AN)**
D(LAKTA(-5))	-1.095741	-0.230301	-0.549012	0.327300
	(0.96055)	(0.17804)	(0.14700)	(0.81387)
	[-1.14074]	[-1.29352]	[-3.73475]	[0.40215]
D(LAKTA(-6))	2.323564	0.178664	0.209795	-1.975258
	(0.90036)	(0.16689)	(0.13779)	(0.76287)
	[2.58069]	[1.07057]	[1.52257]	[-2.58924]
D(LAKTA(-7))	-1.683242	-0.042474	-0.127668	1.712634
	(0.91924)	(0.17038)	(0.14068)	(0.77886)
	[-1.83113]	[-0.24928]	[-0.90752]	[2.19890]
D(LAKTA(-8))	0.915984	-0.121559	0.222906	-1.314028
	(0.62112)	(0.11513)	(0.09505)	(0.52627)
	[1.47474]	[-1.05587]	[2.34504]	[-2.49689]
D(INF_AN(-1))	0.103059	0.014869	0.050319	0.385852
	(0.15199)	(0.02817)	(0.02326)	(0.12878)
	[0.67808]	[0.52781]	[2.16336]	[2.99628]
D(INF_AN(-2))	0.080768	-0.046756	-0.002971	-0.093376
	(0.15519)	(0.02876)	(0.02375)	(0.13149)
	[0.52045]	[-1.62545]	[-0.12510]	[-0.71014]
D(INF_AN(-3))	0.109397	-0.006178	0.089087	0.217075
	(0.14351)	(0.02660)	(0.02196)	(0.12159)
	[0.76232]	[-0.23227]	[4.05642]	[1.78527]

Corrección de errores:	**D(LPIB)**	**D(U)**	**D(LAKTA)**	**D(INF_AN)**
D(INF_AN(-4))	-0.015901	-0.011832	0.000559	-0.369604
	(0.13551)	(0.02512)	(0.02074)	(0.11482)
	[-0.11734]	[-0.47108]	[0.02695]	[-3.21912]
D(INF_AN(-5))	-0.052237	0.003339	0.004460	0.043540
	(0.12014)	(0.02227)	(0.01839)	(0.10179)
	[-0.43480]	[0.14994]	[0.24259]	[0.42772]
D(INF_AN(-6))	0.144258	-0.037624	0.001825	-0.049815
	(0.10666)	(0.01977)	(0.01632)	(0.09038)
	[1.35245]	[-1.90301]	[0.11179]	[-0.55120]
D(INF_AN(-7))	-0.003892	-0.027045	0.045984	0.226512
	(0.11416)	(0.02116)	(0.01747)	(0.09673)
	[-0.03409]	[-1.27805]	[2.63195]	[2.34169]
D(INF_AN(-8))	0.085272	-0.024361	-0.003719	-0.358411
	(0.08847)	(0.01640)	(0.01354)	(0.07496)
	[0.96388]	[-1.48562]	[-0.27472]	[-4.78147]
C	2.175670	-0.280481	0.728128	-0.528191
	(1.02972)	(0.19086)	(0.15759)	(0.87248)
	[2.11287]	[-1.46953]	[4.62049]	[-0.60539]
DD09Q1	-0.050083	0.001400	-0.004012	-0.001302
	(0.01371)	(0.00254)	(0.00210)	(0.01161)
	[-3.65354]	[0.55105]	[-1.91259]	[-0.11212]

Corrección de errores:	**D(LPIB)**	**D(U)**	**D(LAKTA)**	**D(INF_AN)**
DD98Q4	-0.016012	0.003921	-0.001423	0.027784
	(0.01226)	(0.00227)	(0.00188)	(0.01039)
	[-1.30606]	[1.72522]	[-0.75841]	[2.67463]
DD05Q1	-0.024637	0.004963	-0.000795	0.002193
	(0.01013)	(0.00188)	(0.00155)	(0.00858)
	[-2.43300]	[2.64440]	[-0.51318]	[0.25563]
DD98Q1	0.024389	-0.000299	0.008327	0.020066
	(0.01555)	(0.00288)	(0.00238)	(0.01317)
	[1.56857]	[-0.10379]	[3.49947]	[1.52311]
DD99Q4	-0.009659	0.004264	0.002123	-0.016508
	(0.01187)	(0.00220)	(0.00182)	(0.01006)
	[-0.81364]	[1.93796]	[1.16842]	[-1.64129]
DD09Q2	-0.061366	-0.000624	-0.002885	0.001753
	(0.01307)	(0.00242)	(0.00200)	(0.01107)
	[-4.69604]	[-0.25768]	[-1.44252]	[0.15829]

Apéndice C

Código de implementación de red LSTM en Python

```
#Generales
import matplotlib.pyplot as plt
import pandas as pd
import numpy as np
#Preprocesamiento
from sklearn.linear_model import LinearRegression
from sklearn.preprocessing import PolynomialFeatures,MinMaxScaler
from sklearn.pipeline import make_pipeline
from sklearn.metrics import mean_absolute_error
#Redes

from tensorflow.keras.callbacks import EarlyStopping
from sklearn.preprocessing import MinMaxScaler
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import LSTM, Dense, Dropout
from tensorflow.keras.utils import set_random_seed
#Metricas
import statsmodels.api as sm
from statsmodels.tsa.stattools import adfuller
from statsmodels.stats.outliers_influence import variance_inflation_factor

preprocesar=True
set_random_seed(100)
df_original = pd.read_excel('/content/drive/MyDrive/data_final.xlsx')

mejor_mae_LSTM = float('inf')
mejor_modelo_LSTM = None
```

```
mejores_condiciones_LSTM = None

batch = 7
datos=[]

for p_a in range(1, batch):
    df = df_original.copy().drop(["date"], axis=1)
    variables = df.columns
    esc_0=[0,1,2,3]
    esc_1=[0,1,2,4]

    if preprocesar:
        scaler = MinMaxScaler()
        df_scaled = scaler.fit_transform(np.array(df))
        df = pd.DataFrame(df_scaled, columns=variables)

    def crear_secuencias(df, n_steps, target_column):
        X, y = [], []
        for i in range(len(df) - n_steps):
            X.append(df.iloc[i:i + n_steps].values)
            y.append(df.iloc[i + n_steps][target_column])
        return np.array(X), np.array(y)

    n_steps = p_a
    X, y = crear_secuencias(df, n_steps, "lpib")

    test_size=60
    train_size=df.shape[0]-test_size
    X_train, X_test = X[:train_size,:,esc_0], X[train_size:,:,esc_0]
    y_train, y_test = y[:train_size], y[train_size:]
```

```
model = Sequential()
model.add(LSTM(30, activation='tanh',return_sequences=False))
model.add(Dense(1))
model.compile(optimizer='Adam', loss='mse')
model.fit(X_train, y_train, epochs=100, batch_size=32, verbose=0)

pred_train_sc = model.predict(X_train)
pred_test_sc = model.predict(X_test)

ones=np.ones(len(pred_train_sc))
pred_train_sc=np.c_[pred_train_sc,ones,ones,ones,ones,ones,ones]
twos=np.ones(len(pred_test_sc))
pred_test_sc=np.c_[pred_test_sc,twos,twos,twos,twos,twos,twos]

pred_train=scaler.inverse_transform(pred_train_sc)
pred_test=scaler.inverse_transform(pred_test_sc)
pred_train = pred_train[:,0].flatten()
pred_test = pred_test[:,0].flatten()

y_train_sc=np.c_[(y_train.reshape(-1, 1)),ones,ones,ones,ones,ones,ones]
y_test_sc=np.c_[(y_test.reshape(-1, 1)),twos,twos,twos,twos,twos,twos]
y_train=scaler.inverse_transform(y_train_sc)
y_test=scaler.inverse_transform(y_test_sc)
y_train = y_train[:,0].flatten()
y_test = y_test[:,0].flatten()

mae_train = np.mean(np.abs(pred_train - y_train))
mae_test = np.mean(np.abs(pred_test - y_test))
overfitting = abs(mae_test - mae_train)

#guardar resultados en excel y después graficar
```

```
datos.append(pred_train)
datos.append(pred_test)
datos.append(y_train)
datos.append(y_test)

labels=np.array([*range(0,len(y_train),4)])

plt.figure(figsize=(12, 6))
plt.plot(y_train, label='Valores Reales', color='black')
plt.plot(pred_train, label='Predicción LSTM', linestyle=':')
plt.xlabel('Tiempo')
plt.ylabel('ln(PIB)')
plt.title(f'Entrenamiento: Predicciones contra valores reales de ln(PIB) con {p_a} rezagos')
plt.xticks(labels,np.array(df_original['date'][labels]))
plt.tick_params(axis='x', labelrotation = 60, labelsize = 10, colors = 'black')
plt.legend()
plt.show()

plt.figure(figsize=(12, 6))
plt.plot(y_test, label='Valores Reales', color='black')
plt.plot(pred_test, label='Predicción LSTM', linestyle=':')
plt.xlabel('Tiempo')
plt.ylabel('ln(PIB)')
plt.title(f'Prueba: Predicciones contra valores reales de ln(PIB) con {p_a} rezagos')
plt.xticks(labels,np.array(df_original['date'][labels+test_size-1]))
plt.tick_params(axis='x', labelrotation = 60, labelsize = 10, colors = 'black')
plt.legend()
plt.show()

if mae_test < mejor_mae_LSTM:
   mejor_mae_LSTM = mae_test
```

```
        mejor_modelo_LSTM = model
        mejores_condiciones_LSTM = p_a

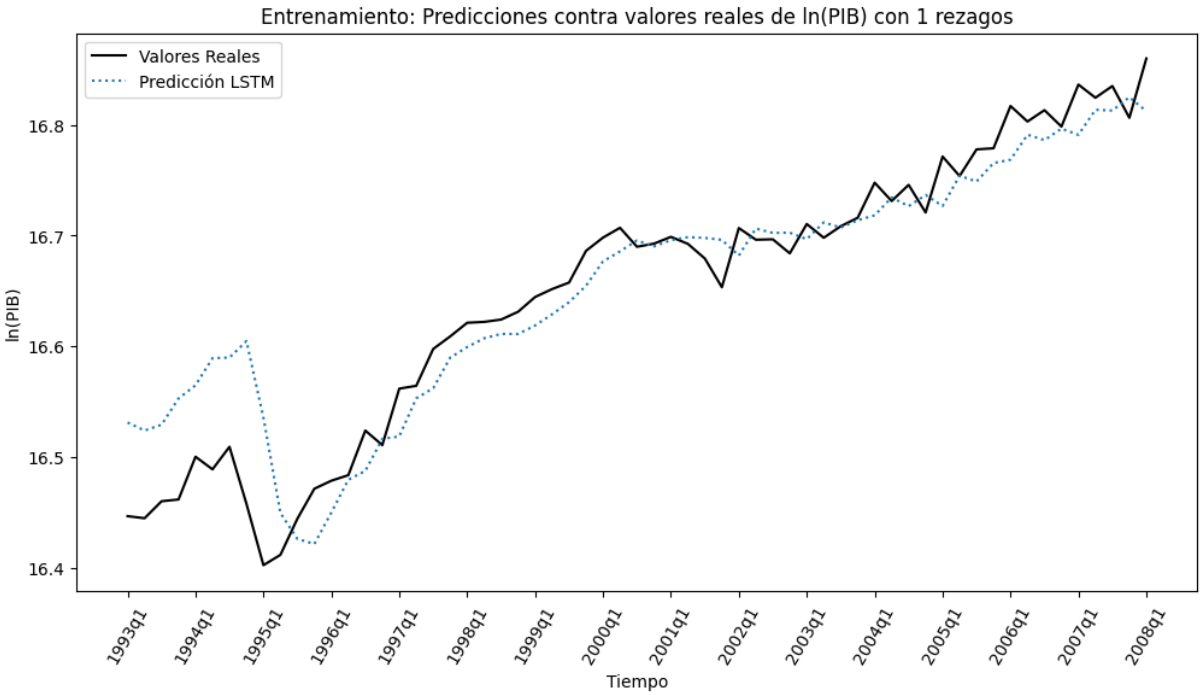
    print(f"Período {p_a}: MAE en train = {mae_train:.4f}")
    print(f"Período {p_a}: MAE en test = {mae_test:.4f}")

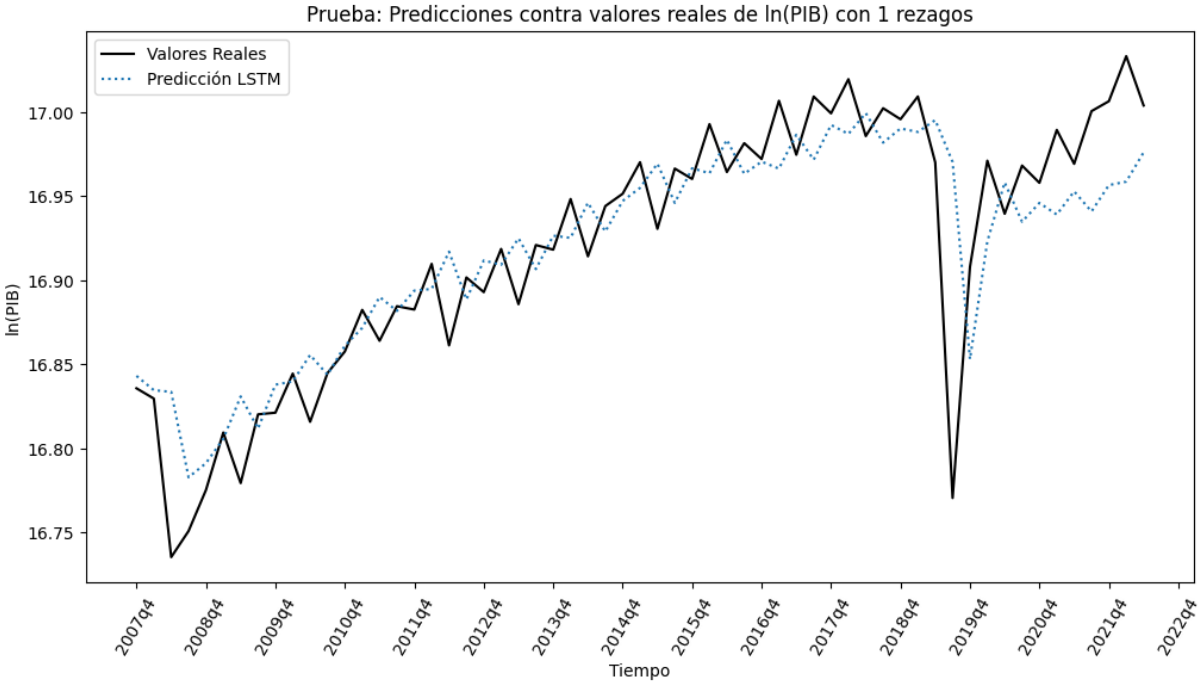
    print(f"Índice de Overfitting : {overfitting}\n")

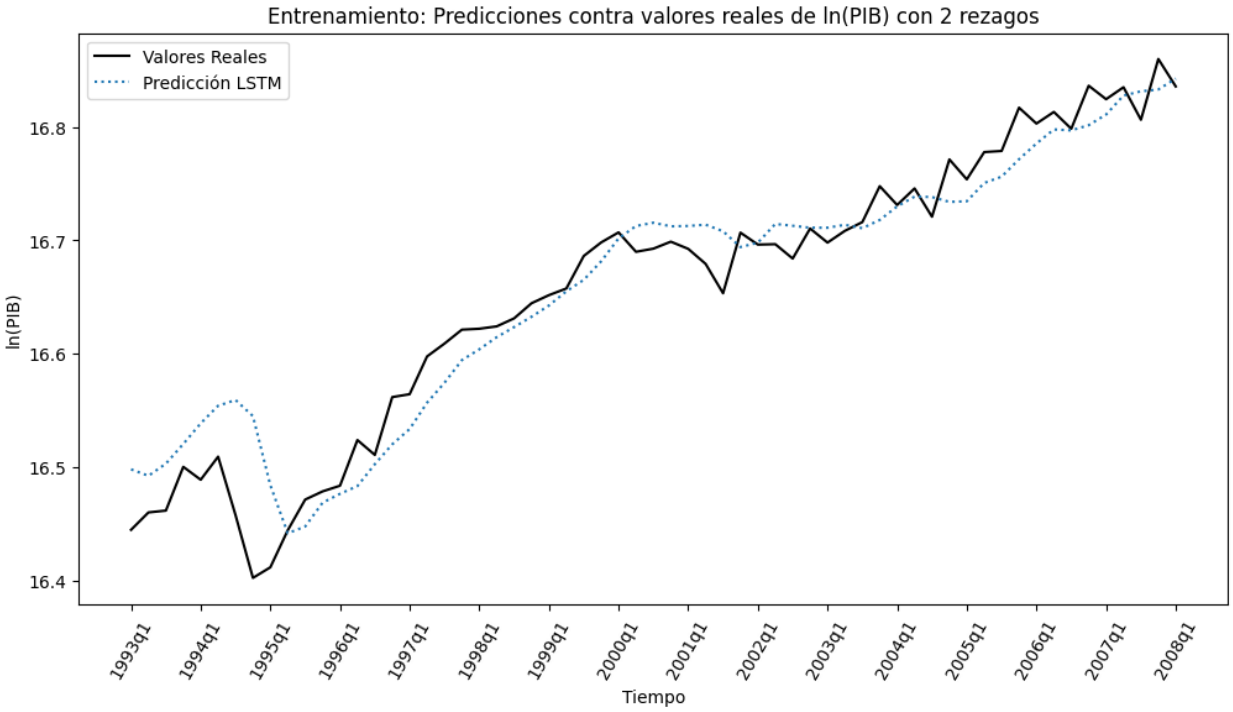
print(f"Mejor MAE: {mejor_mae_LSTM:.4f} con {mejores_condiciones_LSTM} períodos anteriores\n")
datos_df = pd.DataFrame(datos).T
datos_df.to_excel(excel_writer = "/content/drive/MyDrive/data_result_0.xlsx")
print("Resultados guardados en excel")
```

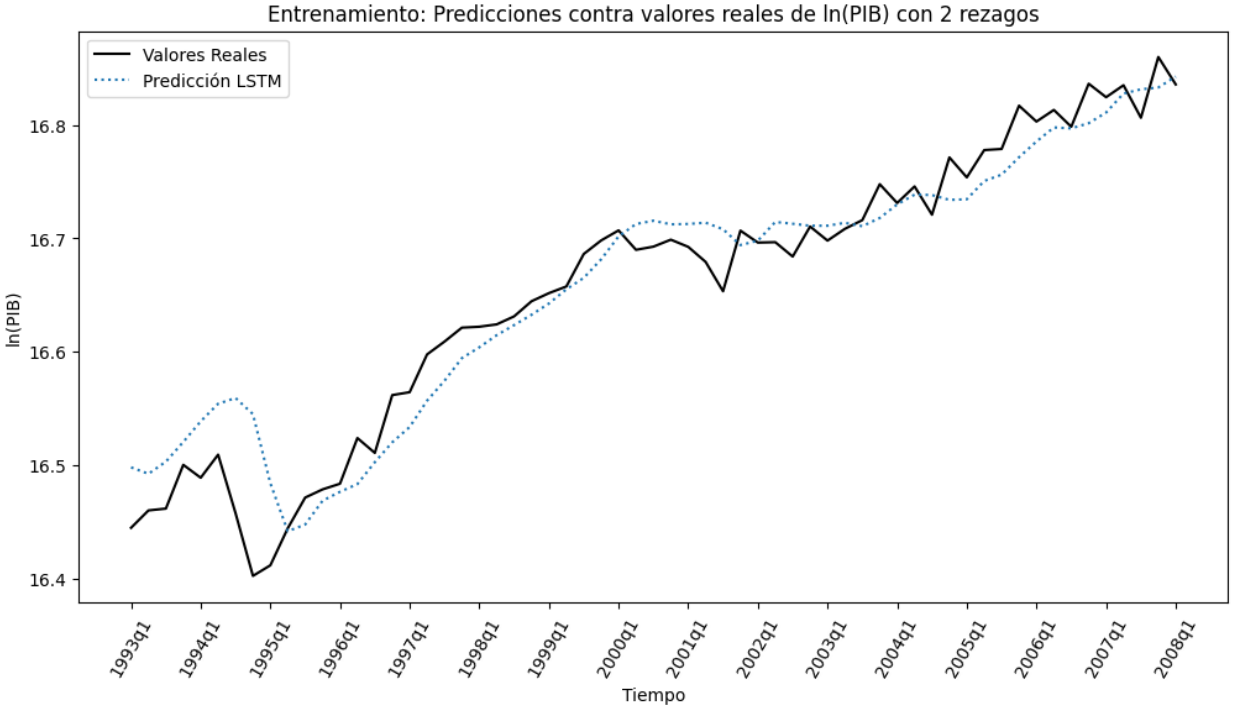
Apéndice D

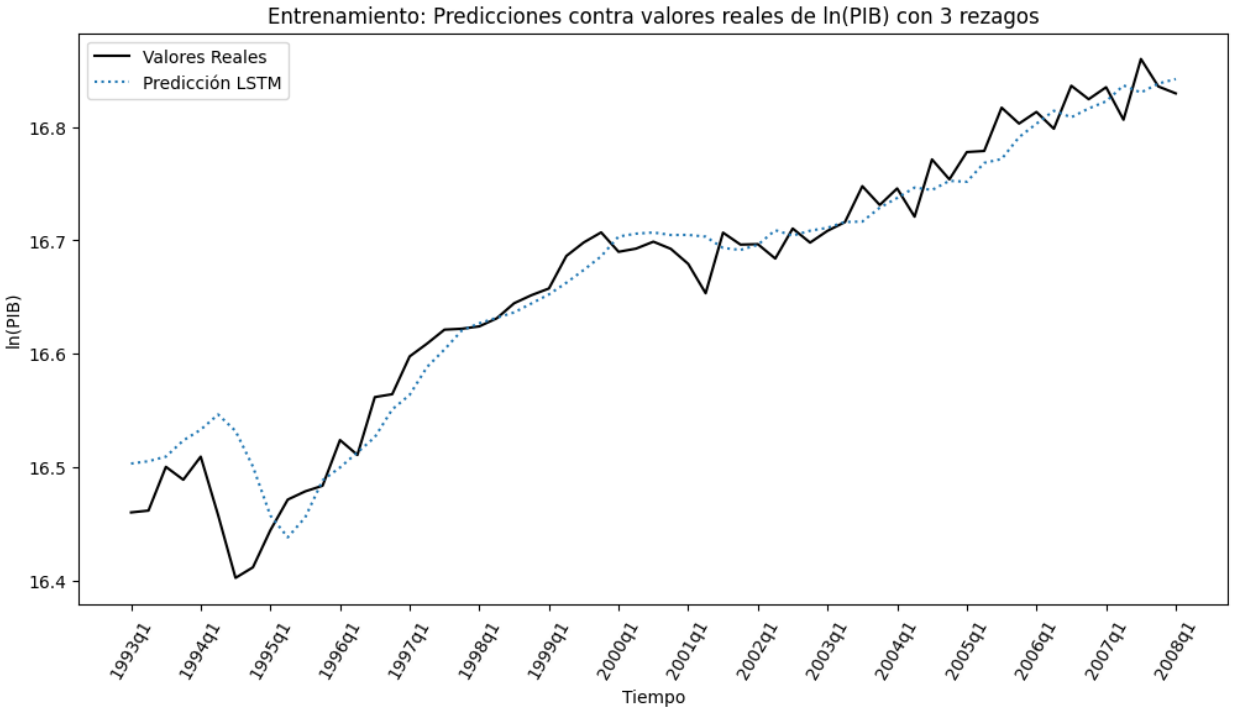
Resultados de entrenamiento y prueba para estimación de ln(PIB) en los rezagos 1,2,3,5,6.

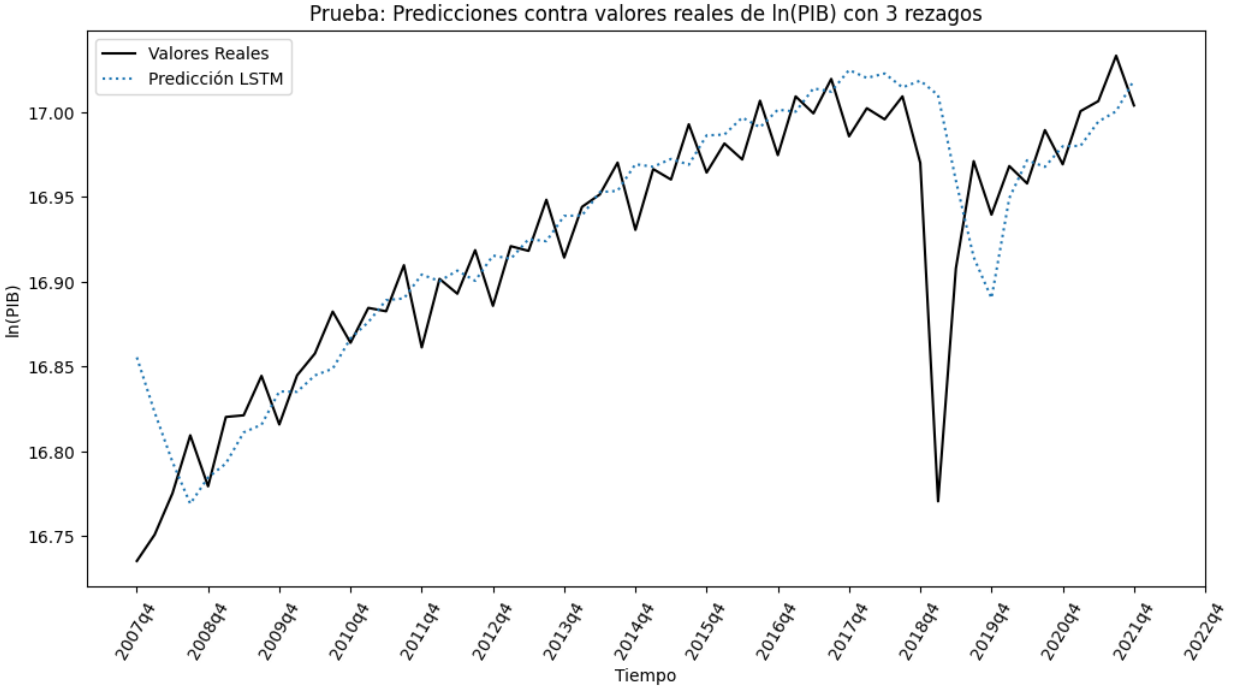


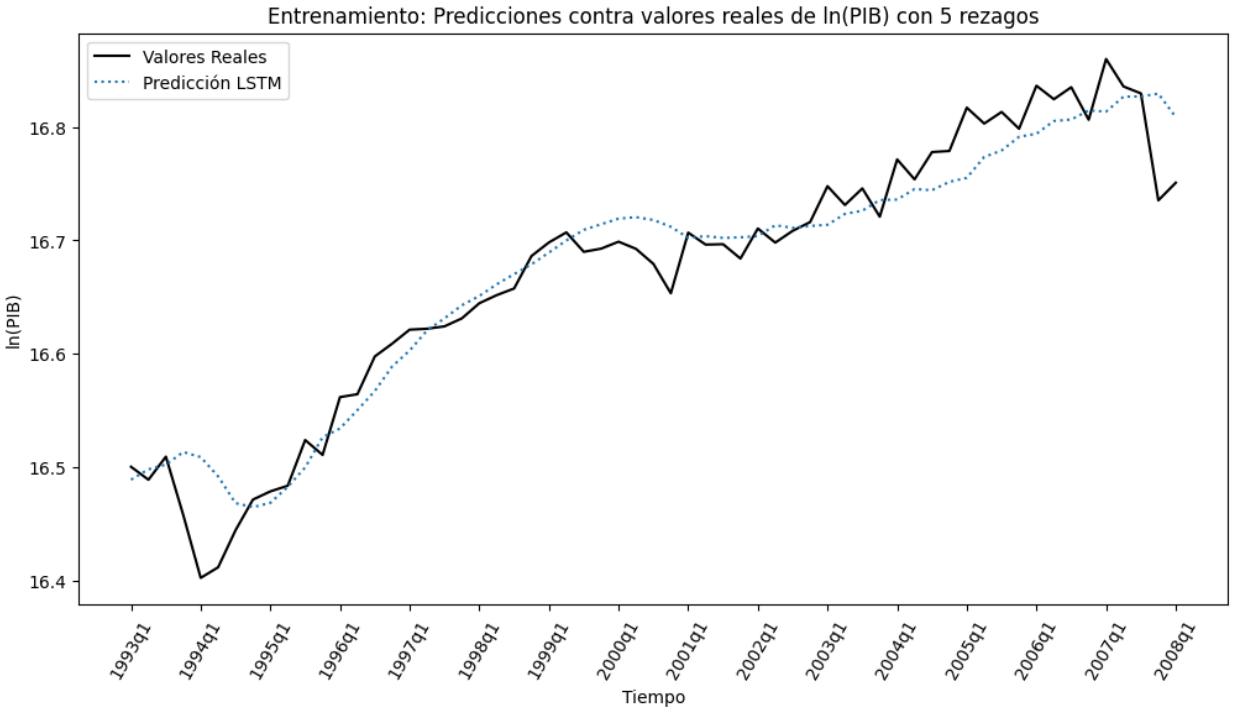


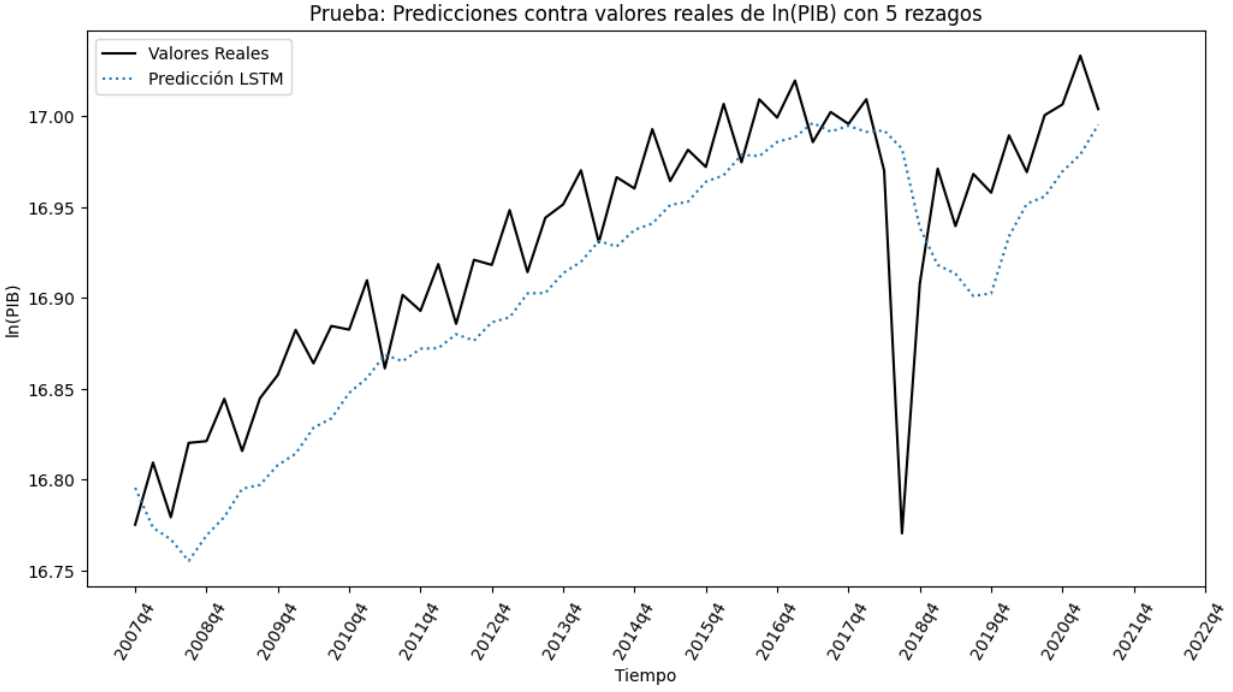

Entrenamiento: Predicciones contra valores reales de ln(PIB) con 2 rezagos
Valores Reales
Predicción LSTM
ln(PIB)
16.8
16.7
16.6
16.5
16.4
1993q1
1994q1
1995q1
1996q1
1997q1
1998q1
1999q1
2000q1
2001q1
2002q1
2003q1
2004q1
2005q1
2006q1
2007q1
2008q1
Tiempo

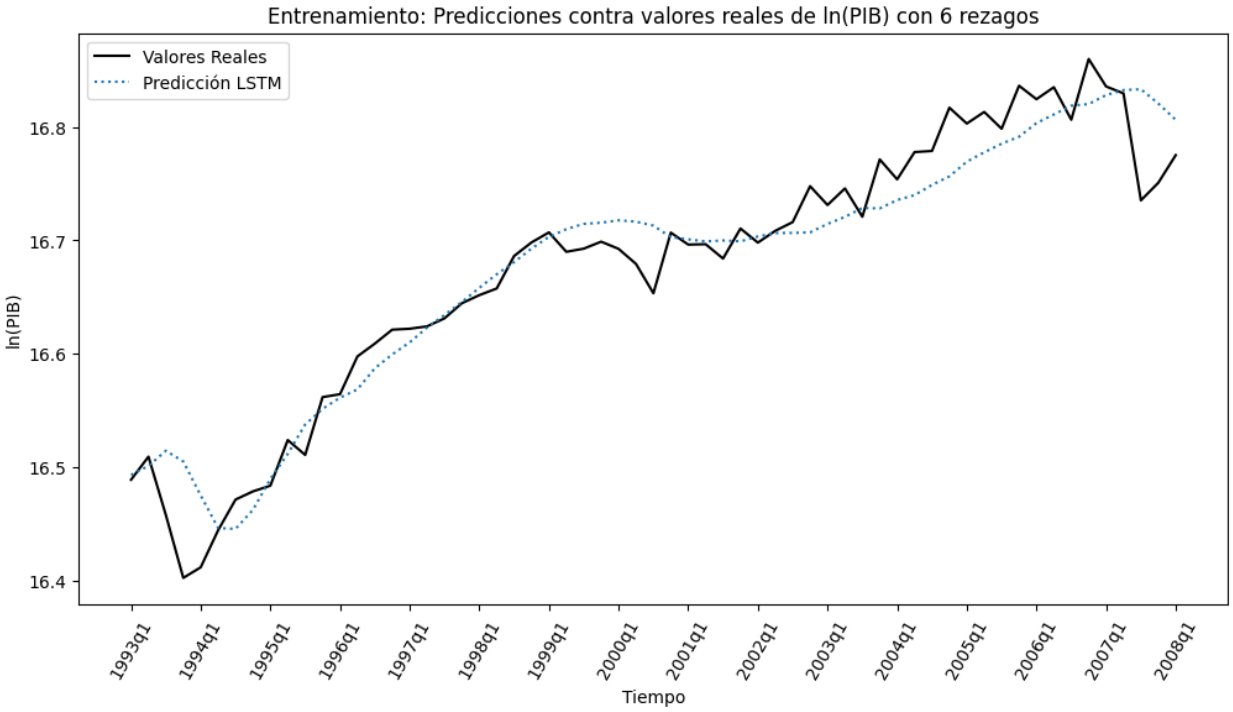

Entrenamiento: Predicciones contra valores reales de ln(PIB) con 2 rezagos
Valores Reales
Predicción LSTM
ln(PIB)
16.8
16.7
16.6
16.5
16.4
1993q1
1994q1
1995q1
1996q1
1997q1
1998q1
1999q1
2000q1
2001q1
2002q1
2003q1
2004q1
2005q1
2006q1
2007q1
2008q1
Tiempo

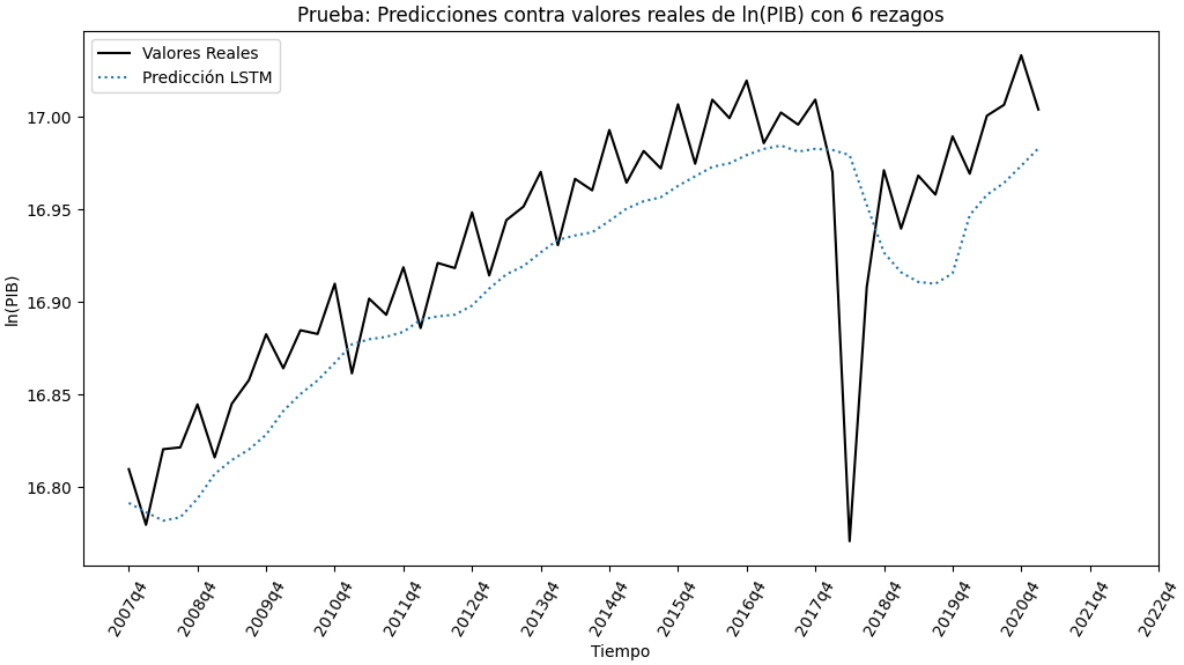

Entrenamiento: Predicciones contra valores reales de ln(PIB) con 3 rezagos
Valores Reales
Predicción LSTM
ln(PIB)
16.8
16.7
16.6
16.5
16.4
1993q1
1994q1
1995q1
1996q1
1997q1
1998q1
1999q1
2000q1
2001q1
2002q1
2003q1
2004q1
2005q1
2006q1
2007q1
2008q1
Tiempo


Prueba: Predicciones contra valores reales de ln(PIB) con 3 rezagos
Valores Reales
Predicción LSTM
ln(PIB)
17.00
16.95
16.90
16.85
16.80
16.75
2007q4
2008q4
2009q4
2010q4
2011q4
2012q4
2013q4
2014q4
2015q4
2016q4
2017q4
2018q4
2019q4
2020q4
2021q4
2022q4
Tiempo

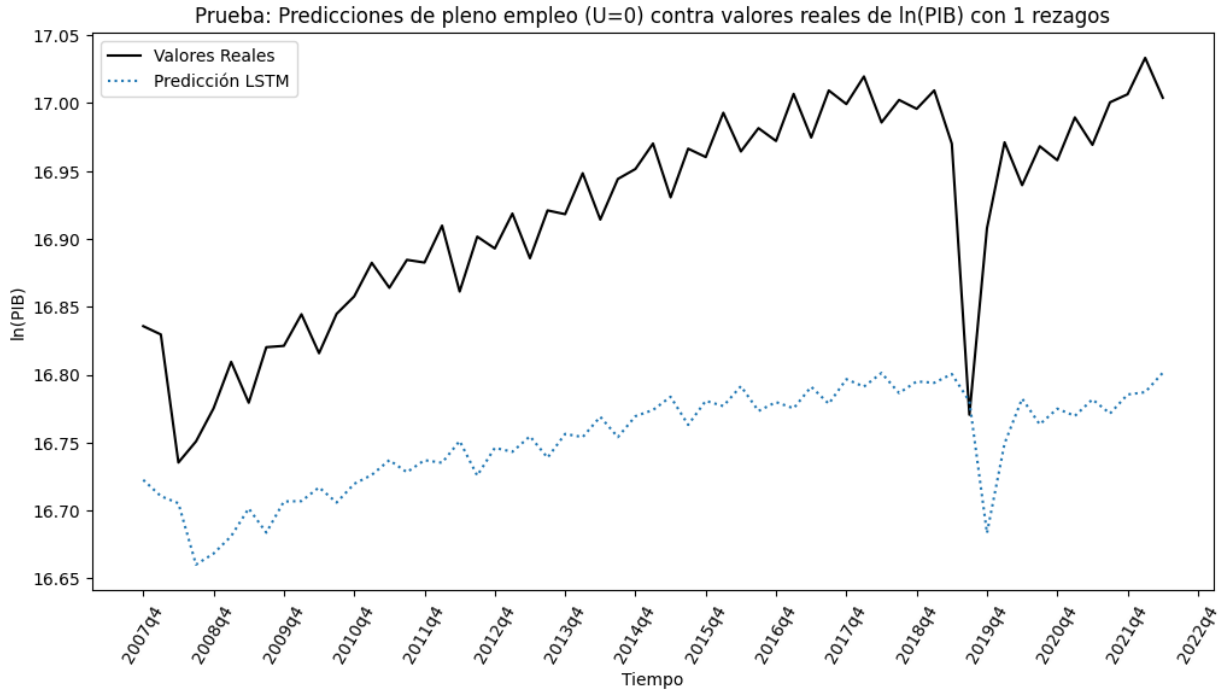

Entrenamiento: Predicciones contra valores reales de ln(PIB) con 5 rezagos
Valores Reales
Predicción LSTM
ln(PIB)
16.8
16.7
16.6
16.5
16.4
1993q1
1994q1
1995q1
1996q1
1997q1
1998q1
1999q1
2000q1
2001q1
2002q1
2003q1
2004q1
2005q1
2006q1
2007q1
2008q1
Tiempo

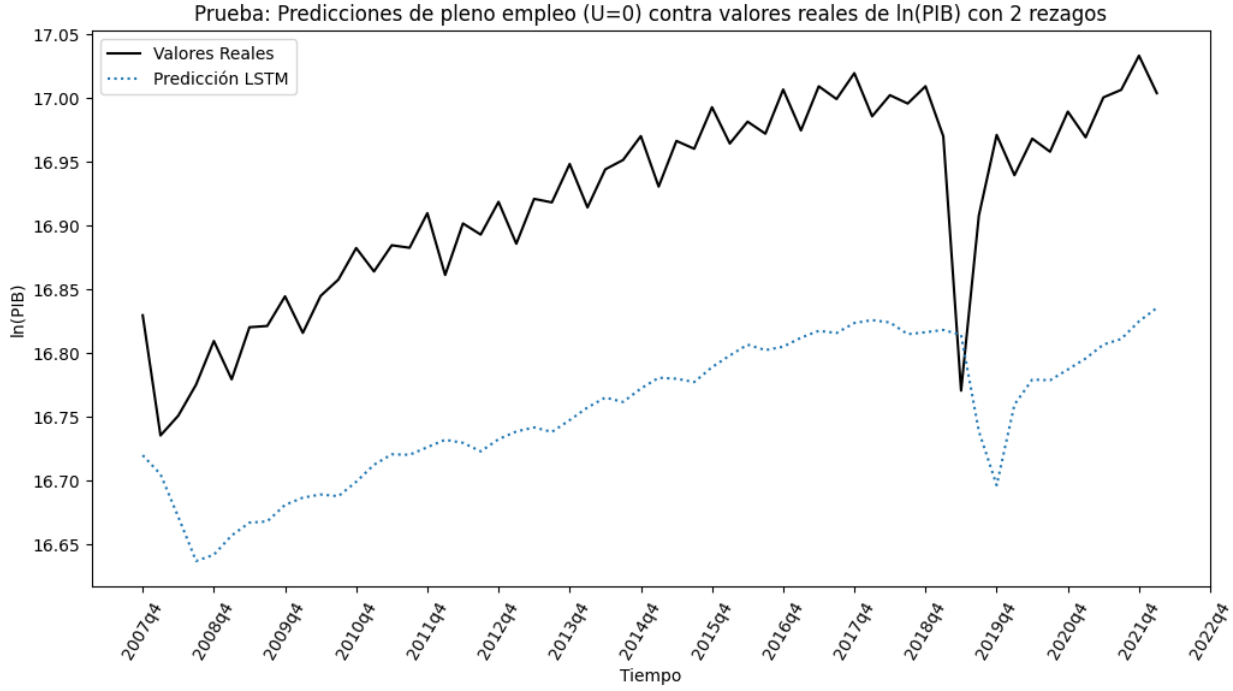

Prueba: Predicciones contra valores reales de ln(PIB) con 5 rezagos
Valores Reales
Predicción LSTM
ln(PIB)
17.00
16.95
16.90
16.85
16.80
16.75
2007q4
2008q4
2009q4
2010q4
2011q4
2012q4
2013q4
2014q4
2015q4
2016q4
2017q4
2018q4
2019q4
2020q4
2021q4
2022q4
Tiempo

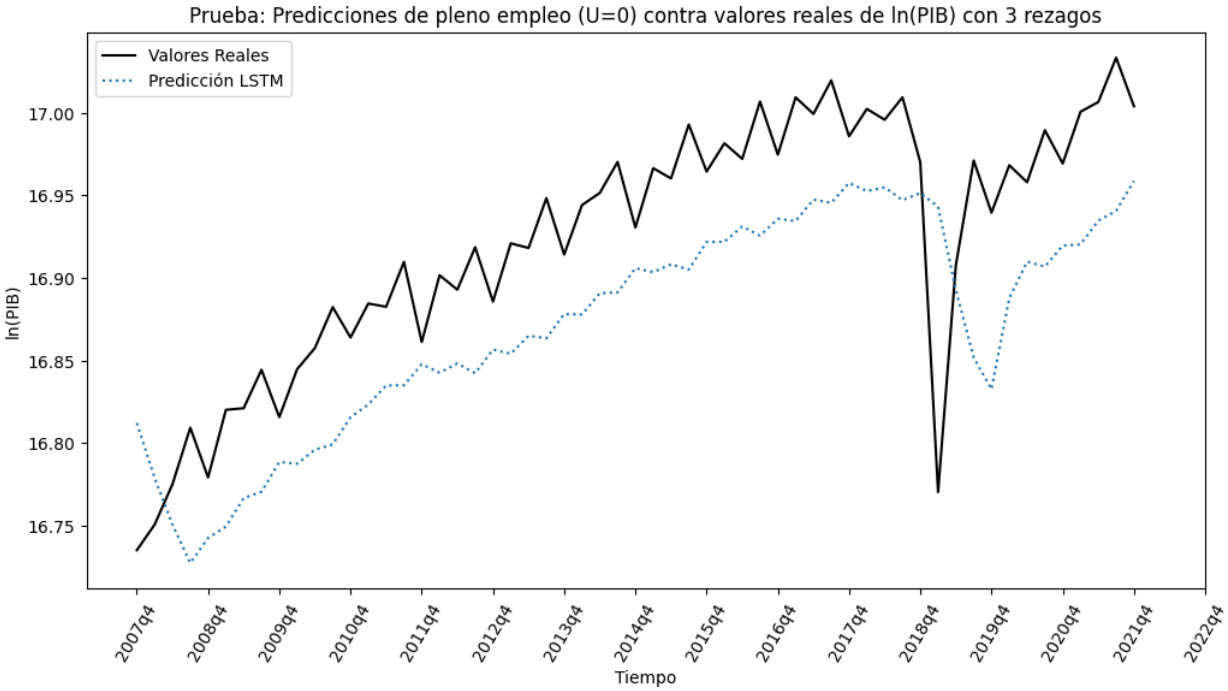

Entrenamiento: Predicciones contra valores reales de ln(PIB) con 6 rezagos
Valores Reales
Predicción LSTM
ln(PIB)
16.8
16.7
16.6
16.5
16.4
1993q1
1994q1
1995q1
1996q1
1997q1
1998q1
1999q1
2000q1
2001q1
2002q1
2003q1
2004q1
2005q1
2006q1
2007q1
2008q1
Tiempo

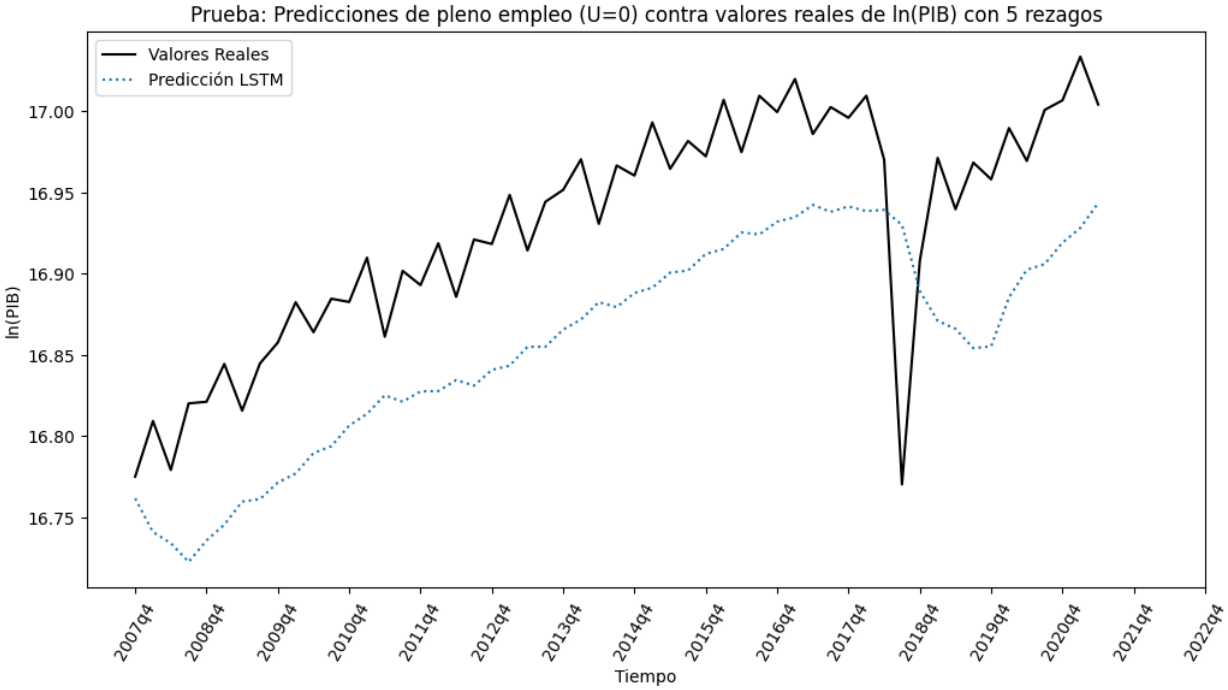

Prueba: Predicciones contra valores reales de ln(PIB) con 6 rezagos
Valores Reales
Predicción LSTM
ln(PIB)
17.00
16.95
16.90
16.85
16.80
2007q4
2008q4
2009q4
2010q4
2011q4
2012q4
2013q4
2014q4
2015q4
2016q4
2017q4
2018q4
2019q4
2020q4
2021q4
2022q4
Tiempo

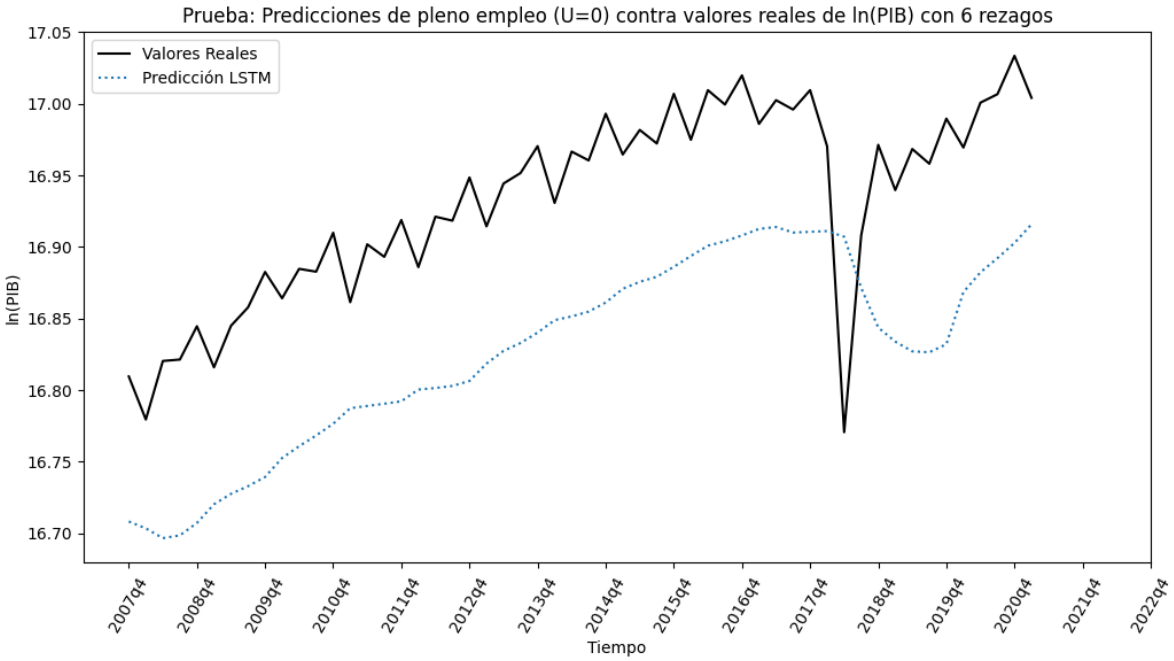
Resultados de prueba para estimación de ln(PIB) potencial en escenario de pleno empleo (u=0) en los rezagos 1,2,3,5,6.



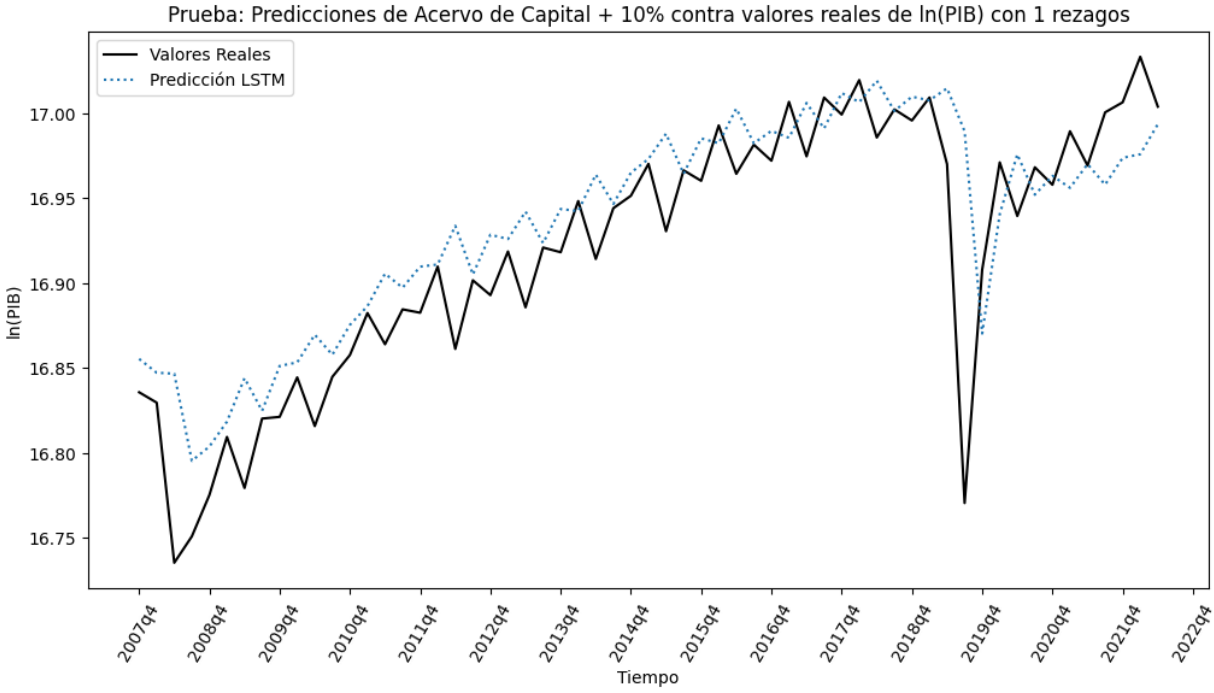


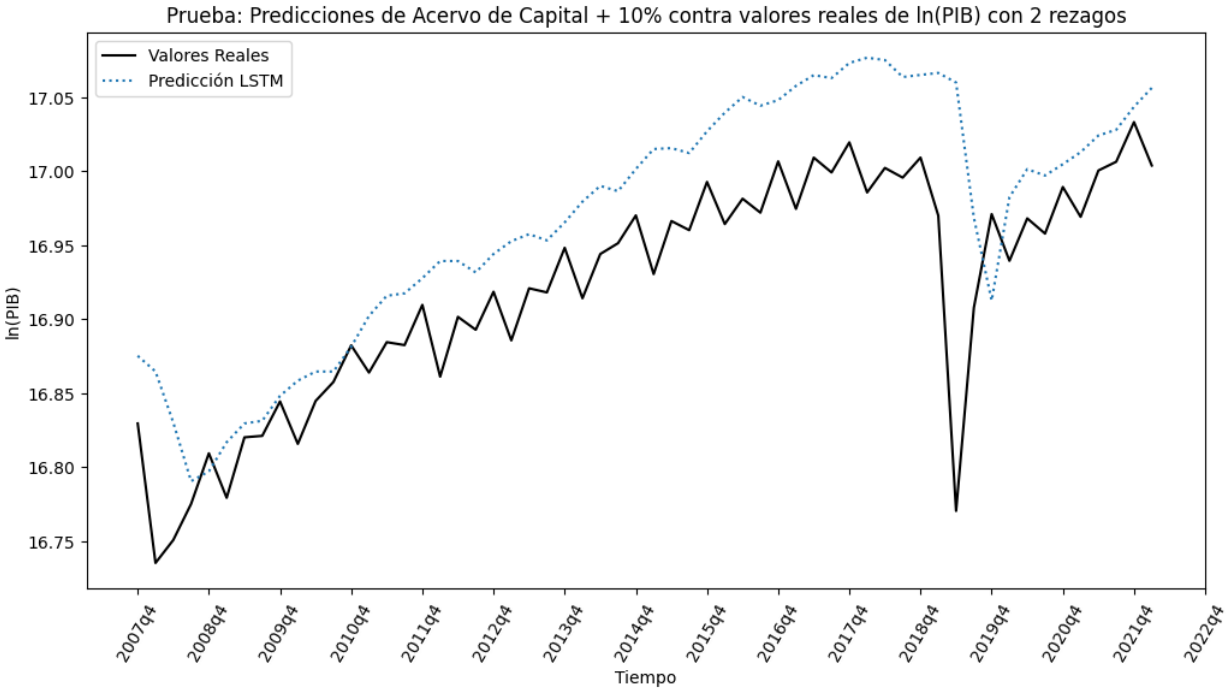

Prueba: Predicciones de pleno empleo (U=0) contra valores reales de ln(PIB) con 3 rezagos
Valores Reales
Predicción LSTM
ln(PIB)
17.00
16.95
16.90
16.85
16.80
16.75
2007q4
2008q4
2009q4
2010q4
2011q4
2012q4
2013q4
2014q4
2015q4
2016q4
2017q4
2018q4
2019q4
2020q4
2021q4
2022q4
Tiempo

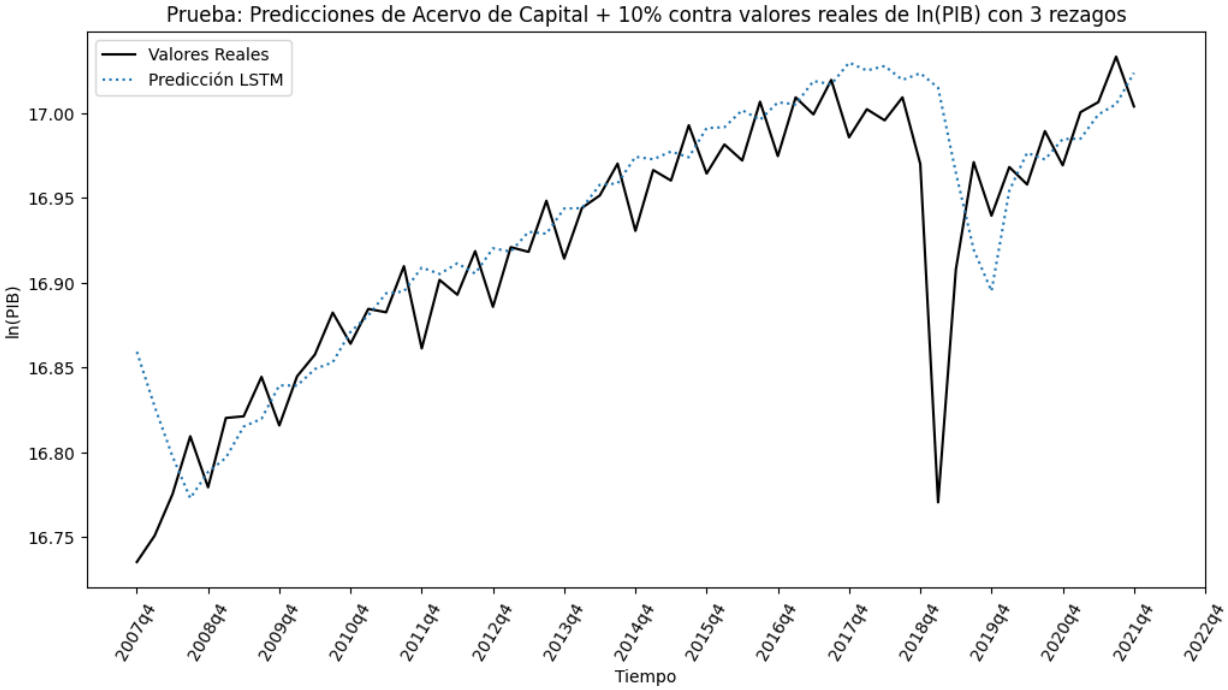

Prueba: Predicciones de pleno empleo (U=0) contra valores reales de ln(PIB) con 5 rezagos
Valores Reales
Predicción LSTM
ln(PIB)
17.00
16.95
16.90
16.85
16.80
16.75
2007q4
2008q4
2009q4
2010q4
2011q4
2012q4
2013q4
2014q4
2015q4
2016q4
2017q4
2018q4
2019q4
2020q4
2021q4
2022q4
Tiempo

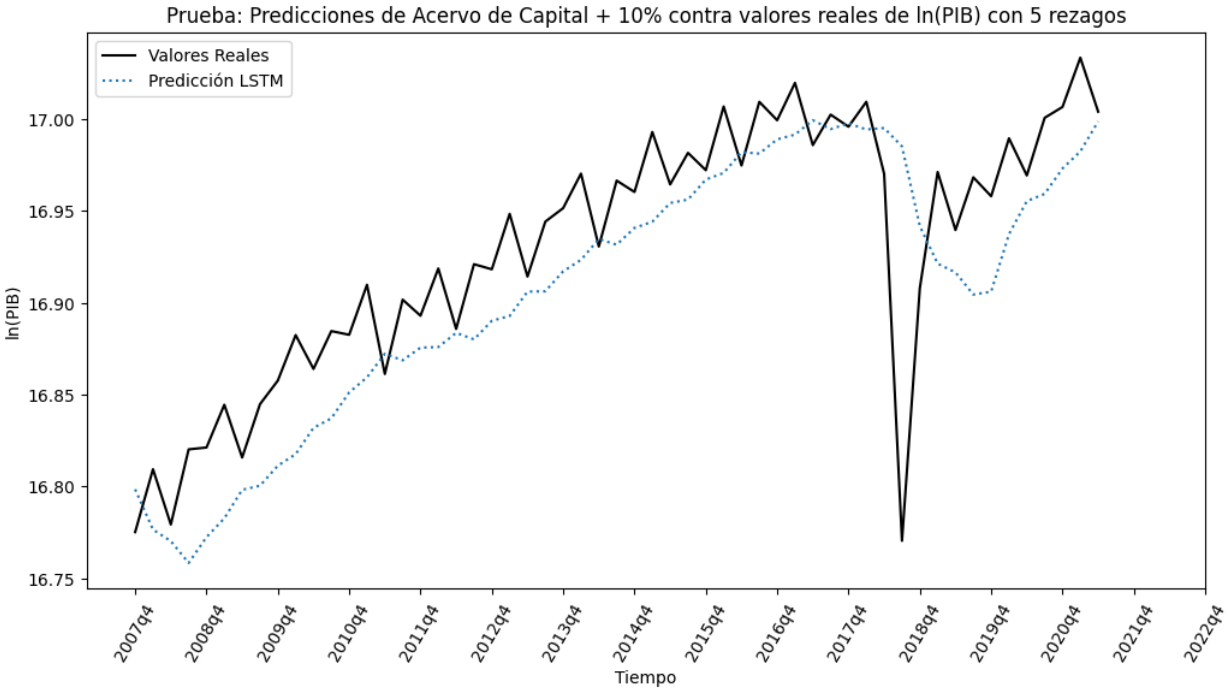


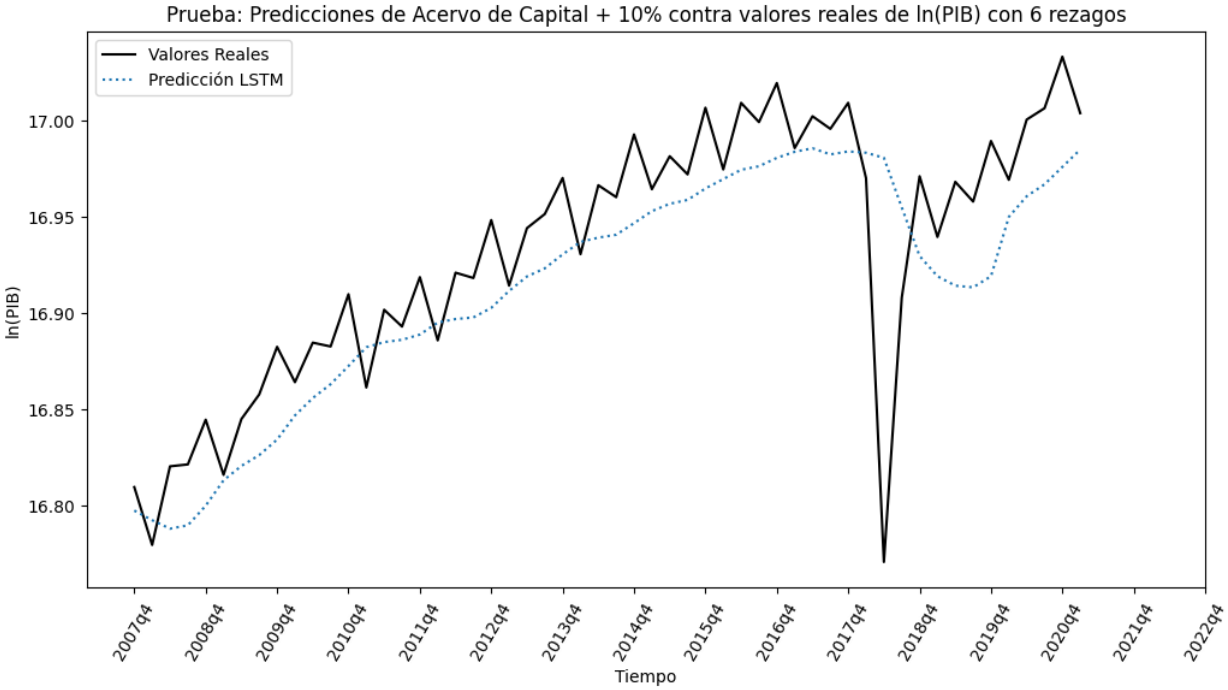
Resultados de prueba para estimación de ln(PIB) potencial en escenario de Acervo de Capital +10% en los rezagos 1,2,3,5,6.



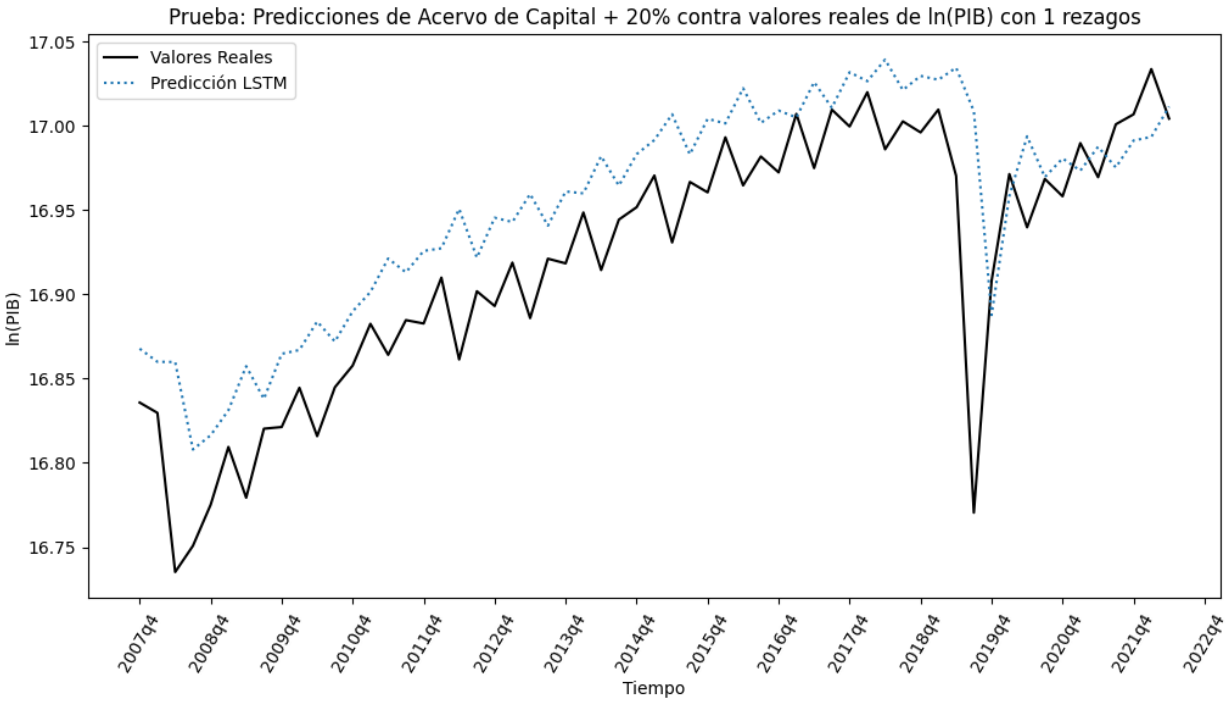

Prueba: Predicciones de Acervo de Capital + 10% contra valores reales de ln(PIB) con 2 rezagos
Valores Reales
Predicción LSTM
ln(PIB)
17.05
17.00
16.95
16.90
16.85
16.80
16.75
2007q4
2008q4
2009q4
2010q4
2011q4
2012q4
2013q4
2014q4
2015q4
2016q4
2017q4
2018q4
2019q4
2020q4
2021q4
2022q4
Tiempo

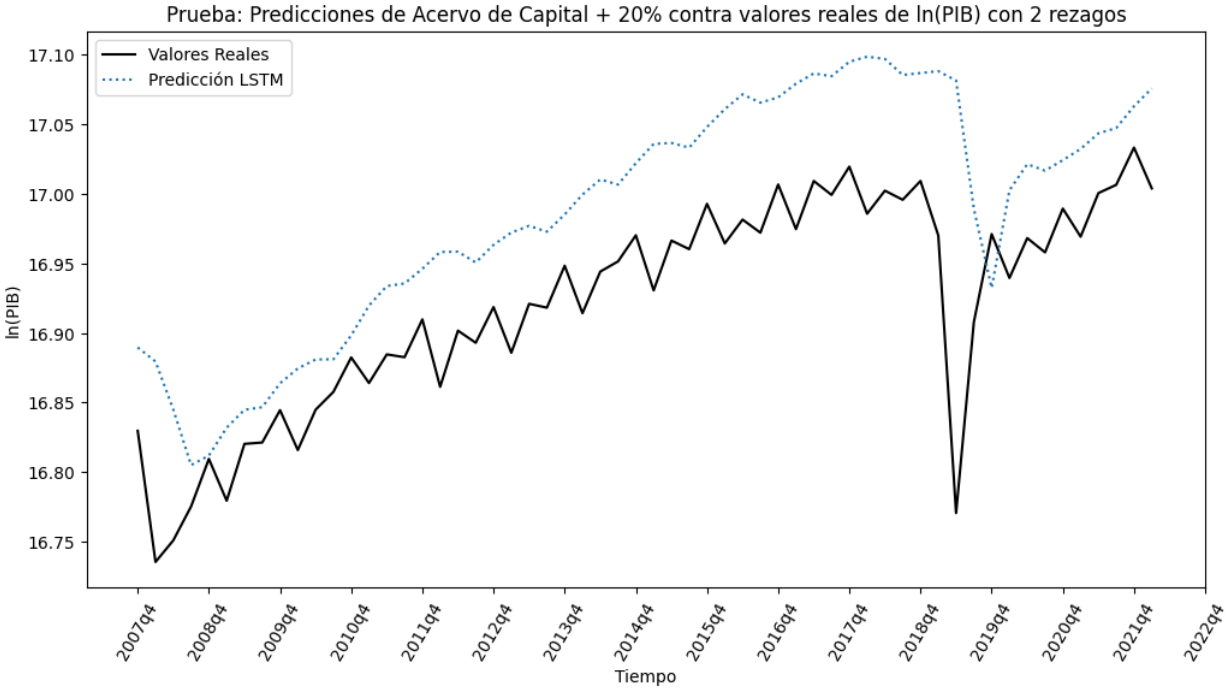

Prueba: Predicciones de Acervo de Capital + 10% contra valores reales de ln(PIB) con 3 rezagos
Valores Reales
Predicción LSTM
ln(PIB)
17.00
16.95
16.90
16.85
16.80
16.75
2007q4
2008q4
2009q4
2010q4
2011q4
2012q4
2013q4
2014q4
2015q4
2016q4
2017q4
2018q4
2019q4
2020q4
2021q4
2022q4
Tiempo

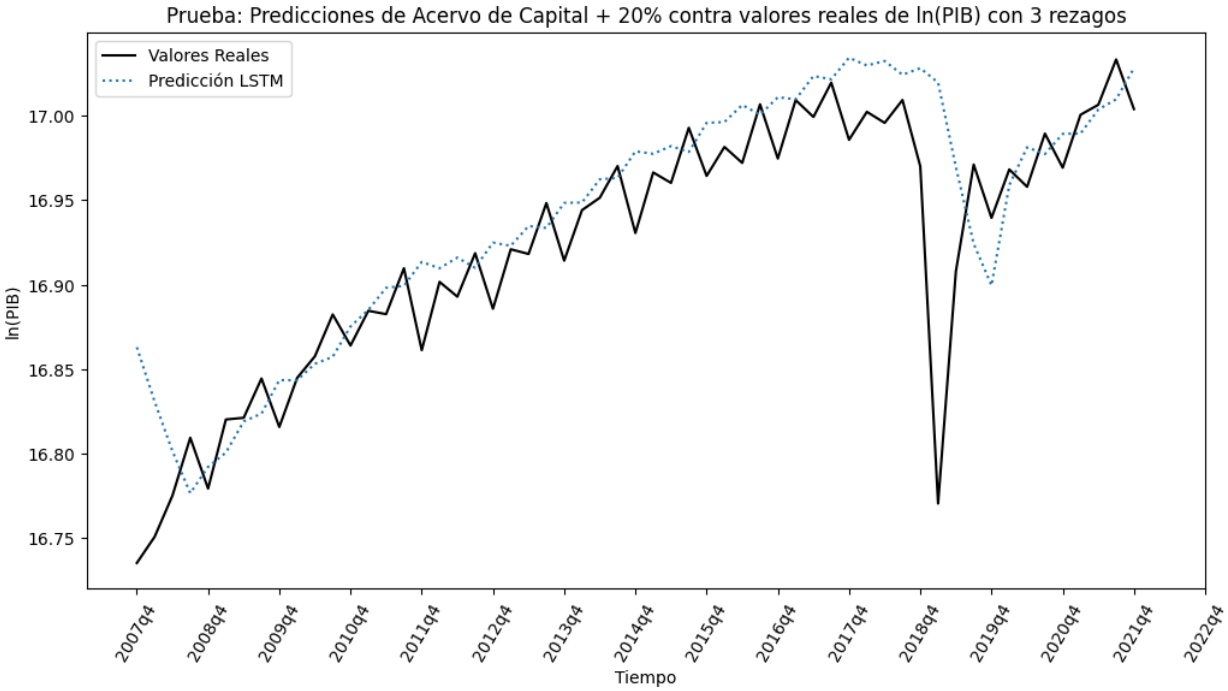

Prueba: Predicciones de Acervo de Capital + 10% contra valores reales de ln(PIB) con 5 rezagos
Valores Reales
Predicción LSTM
ln(PIB)
Tiempo

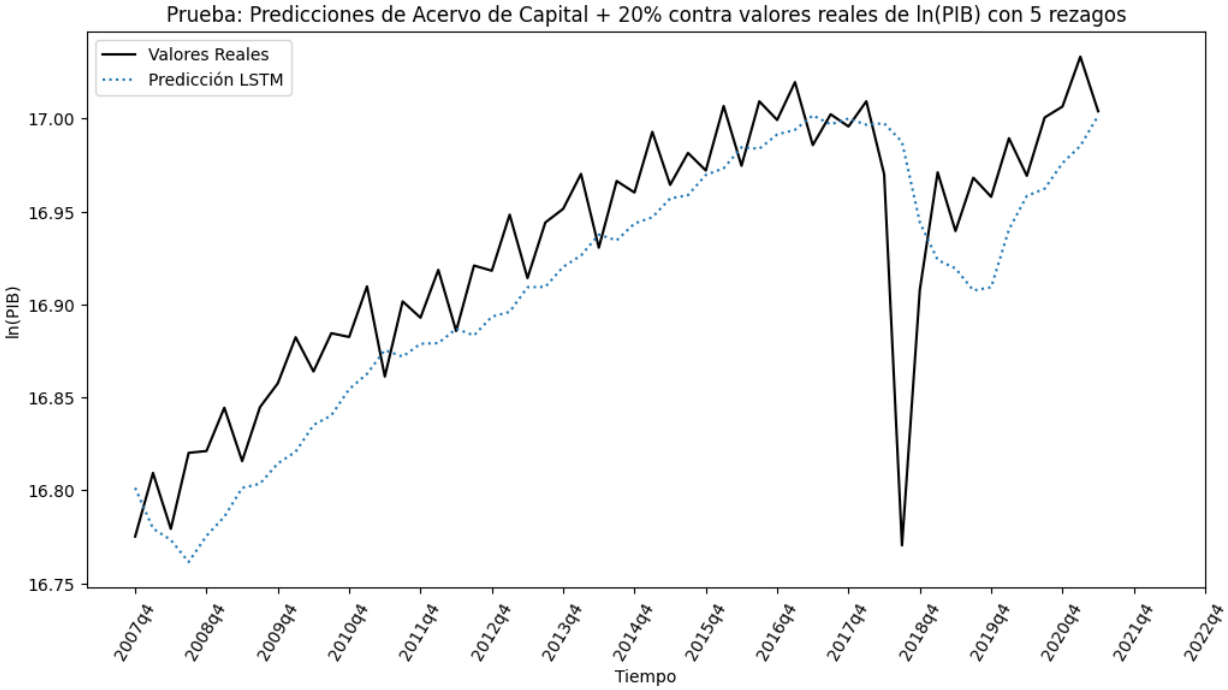

Prueba: Predicciones de Acervo de Capital + 10% contra valores reales de ln(PIB) con 6 rezagos
Valores Reales
Predicción LSTM
ln(PIB)
Tiempo

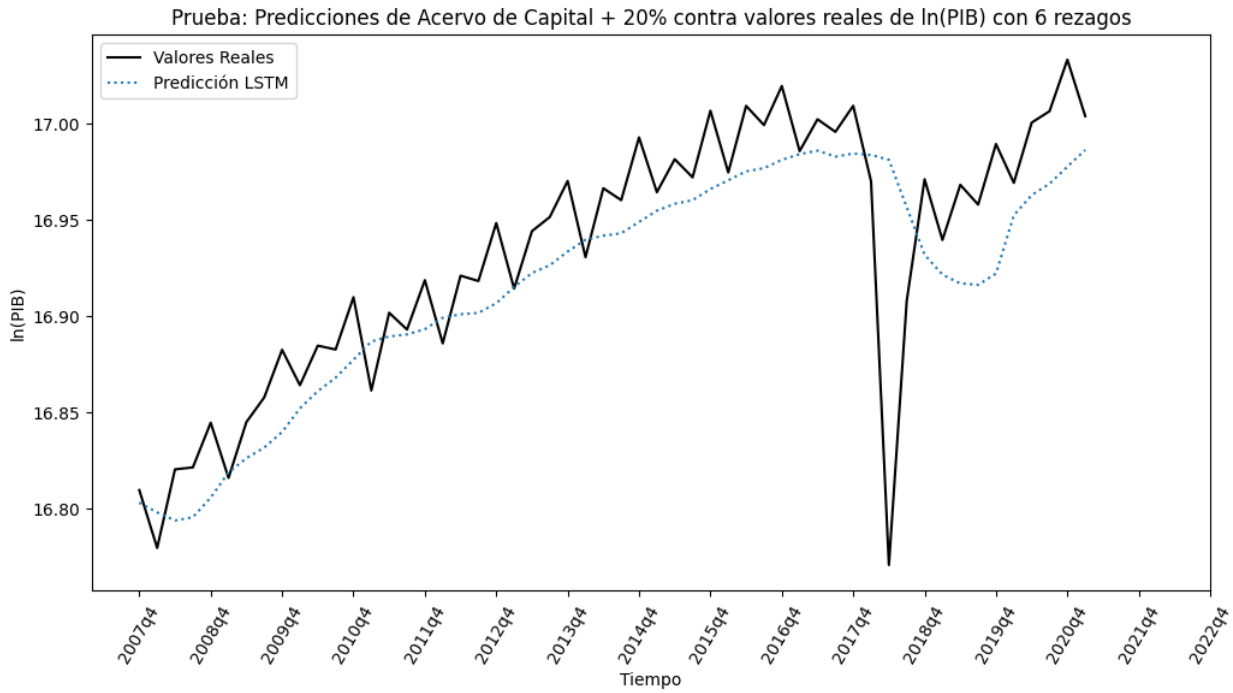
Resultados de prueba para estimación de ln(PIB) potencial en escenario de Acervo de Capital +20% en los rezagos 1,2,3,5,6.



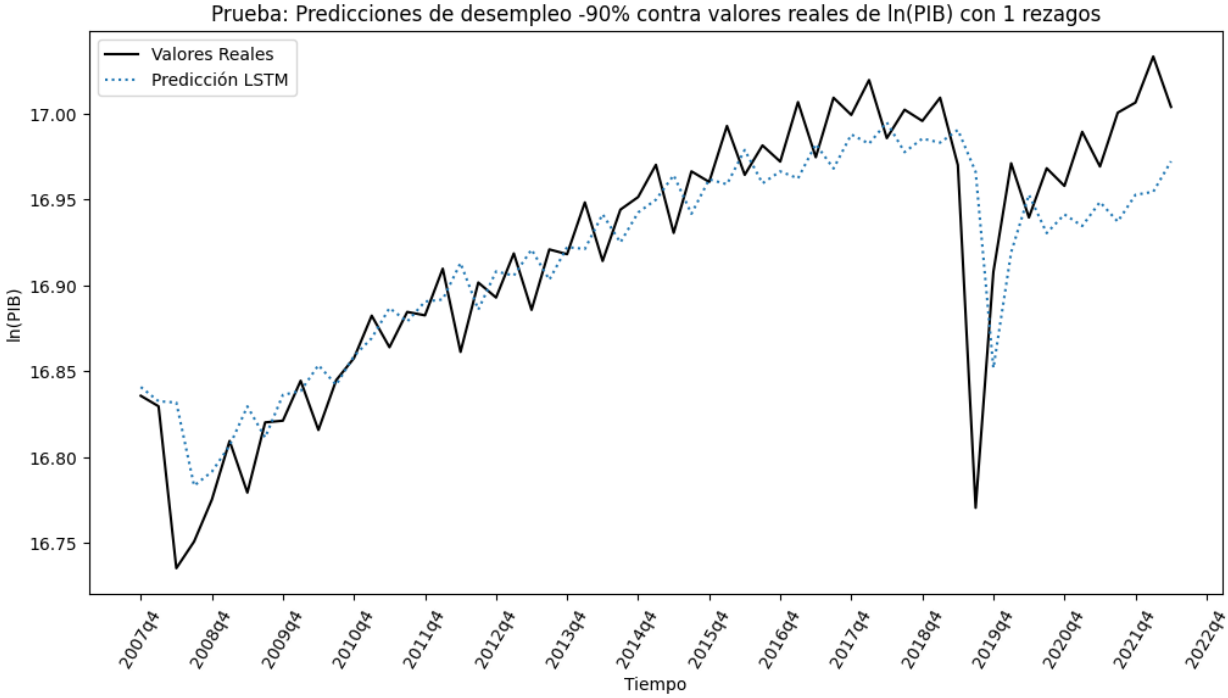


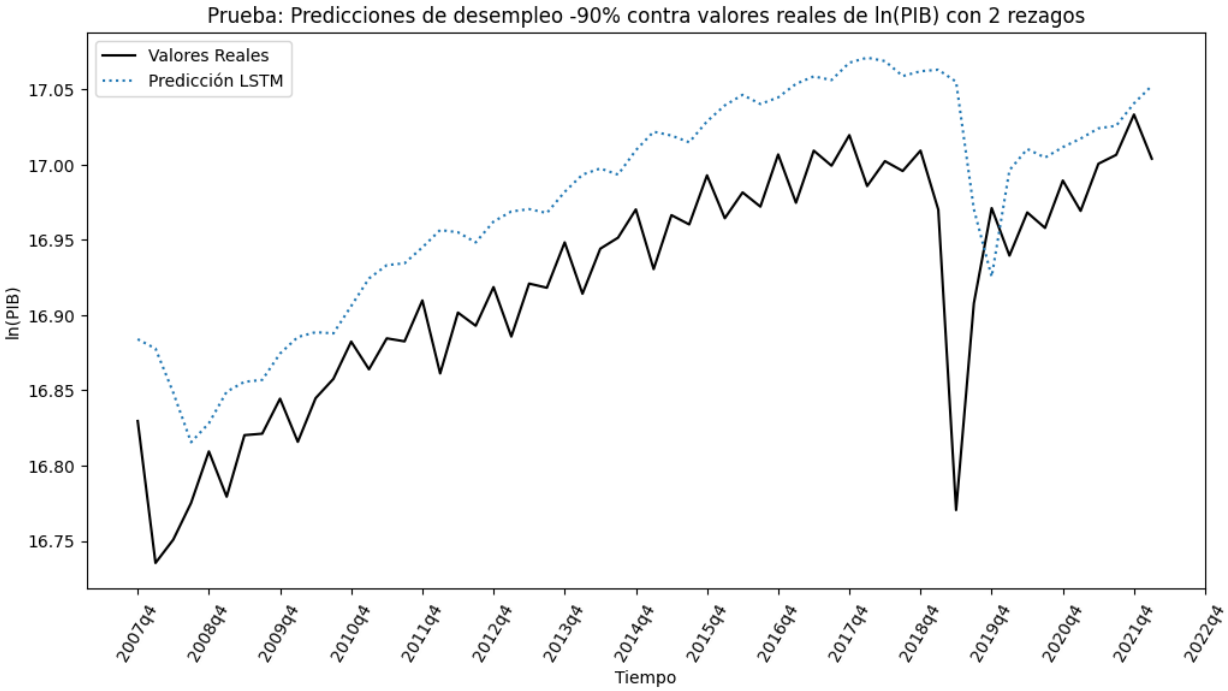

Prueba: Predicciones de Acervo de Capital + 20% contra valores reales de ln(PIB) con 3 rezagos
Valores Reales
Predicción LSTM
ln(PIB)
17.00
16.95
16.90
16.85
16.80
16.75
2007q4
2008q4
2009q4
2010q4
2011q4
2012q4
2013q4
2014q4
2015q4
2016q4
2017q4
2018q4
2019q4
2020q4
2021q4
2022q4
Tiempo

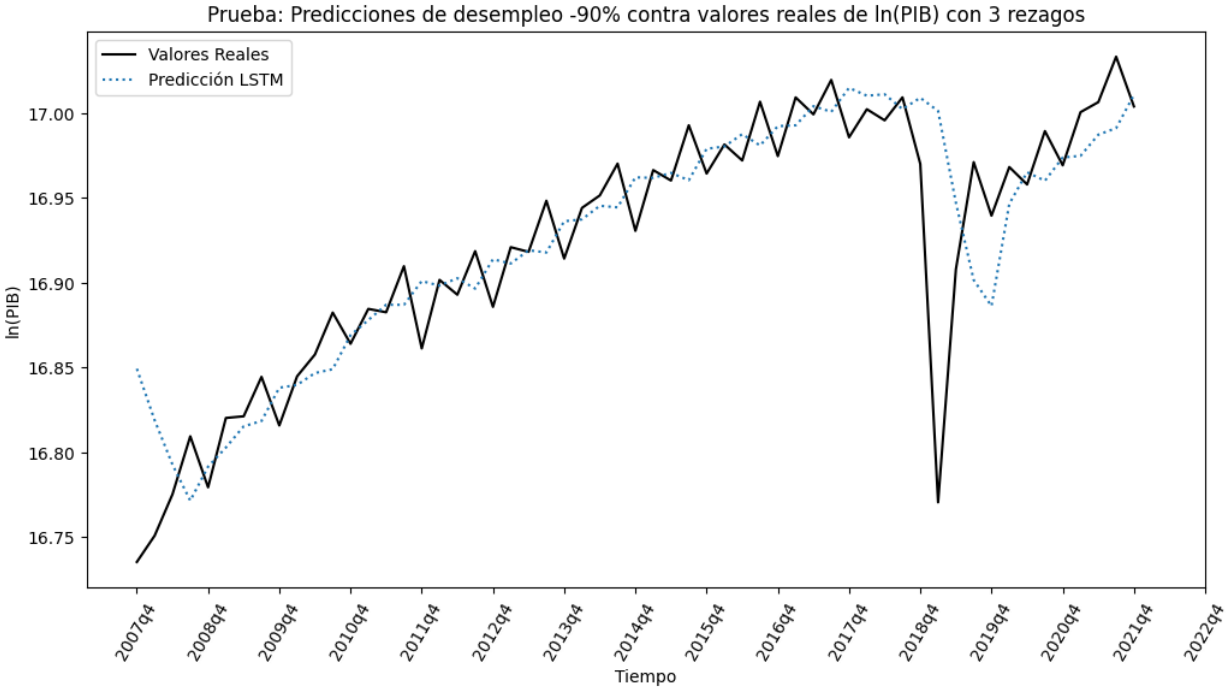

Prueba: Predicciones de Acervo de Capital + 20% contra valores reales de ln(PIB) con 5 rezagos
Valores Reales
Predicción LSTM
ln(PIB)
17.00
16.95
16.90
16.85
16.80
16.75
2007q4
2008q4
2009q4
2010q4
2011q4
2012q4
2013q4
2014q4
2015q4
2016q4
2017q4
2018q4
2019q4
2020q4
2021q4
2022q4
Tiempo

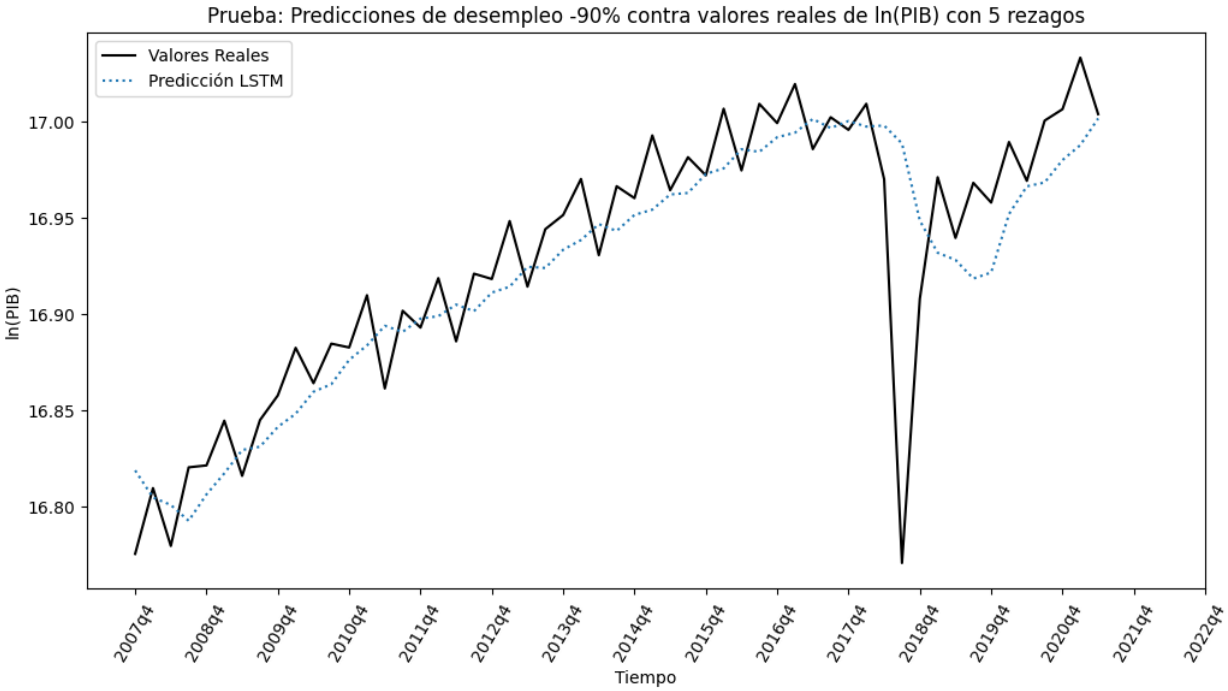


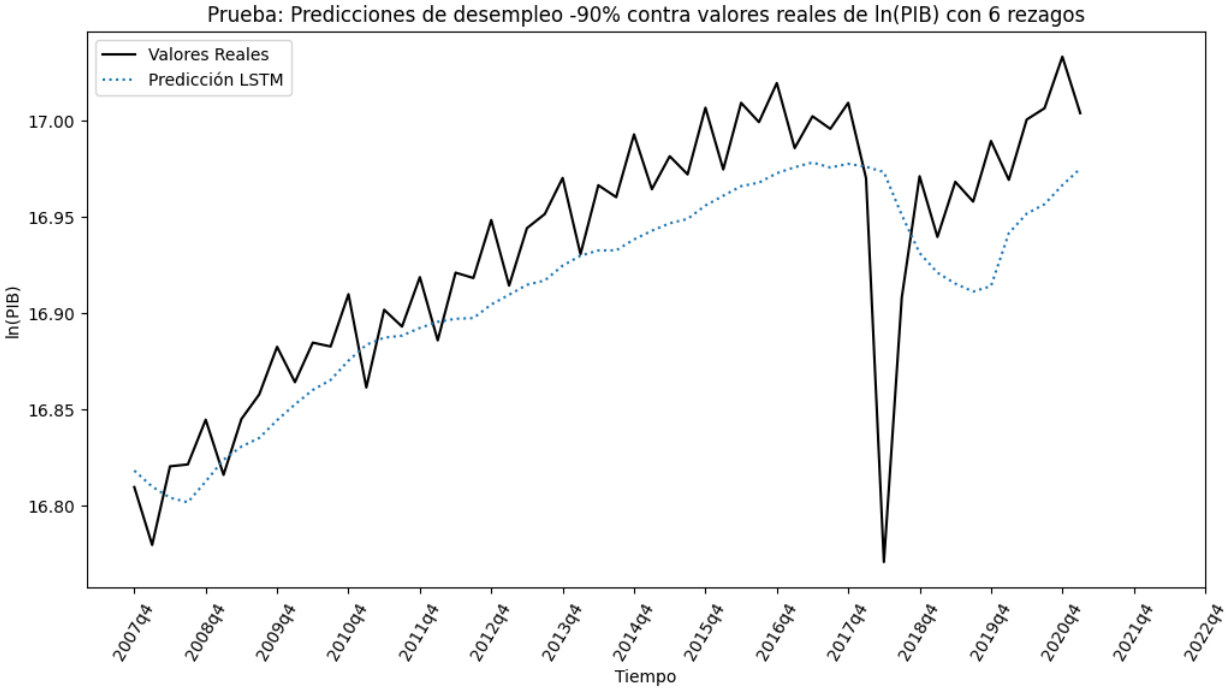
Resultados de prueba para estimación de ln(PIB) potencial en escenario de desempleo -90% en los rezagos 1,2,3,5,6.




Prueba: Predicciones de desempleo -90% contra valores reales de ln(PIB) con 2 rezagos
Valores Reales
Predicción LSTM
ln(PIB)
17.05
17.00
16.95
16.90
16.85
16.80
16.75
2007q4
2008q4
2009q4
2010q4
2011q4
2012q4
2013q4
2014q4
2015q4
2016q4
2017q4
2018q4
2019q4
2020q4
2021q4
2022q4
Tiempo


Prueba: Predicciones de desempleo -90% contra valores reales de ln(PIB) con 3 rezagos
Valores Reales
Predicción LSTM
ln(PIB)
17.00
16.95
16.90
16.85
16.80
16.75
2007q4
2008q4
2009q4
2010q4
2011q4
2012q4
2013q4
2014q4
2015q4
2016q4
2017q4
2018q4
2019q4
2020q4
2021q4
2022q4
Tiempo


Prueba: Predicciones de desempleo -90% contra valores reales de ln(PIB) con 5 rezagos
Valores Reales
Predicción LSTM
ln(PIB)
17.00
16.95
16.90
16.85
16.80
2007q4
2008q4
2009q4
2010q4
2011q4
2012q4
2013q4
2014q4
2015q4
2016q4
2017q4
2018q4
2019q4
2020q4
2021q4
2022q4
Tiempo


Prueba: Predicciones de desempleo -90% contra valores reales de ln(PIB) con 6 rezagos
Valores Reales
Predicción LSTM
ln(PIB)
17.00
16.95
16.90
16.85
16.80
2007q4
2008q4
2009q4
2010q4
2011q4
2012q4
2013q4
2014q4
2015q4
2016q4
2017q4
2018q4
2019q4
2020q4
2021q4
2022q4
Tiempo