

Pronóstico de la inflación en Costa Rica mediante métodos de aprendizaje estadístico

Adolfo Rodríguez Vargas

Diciembre 2019

Jornadas de Investigación Económica 2019



Las ideas aquí expresadas son del autor y no necesariamente representan las del Banco Central de Costa Rica

1- Motivación

- Pronóstico de inflación en BCCR:
 - Conjunto de modelos univariados.
 - Conjunto de modelos bayesianos.
 - Modelo macroeconómico semiestructural.
- Métodos de aprendizaje estadístico (AE) han ganado popularidad como herramienta para el pronóstico.
- La mayoría de aplicaciones usa datos de corte transversal en problemas de clasificación.
- Sin embargo, los métodos de AE también pueden adaptarse para la predicción de variables continuas en un contexto de series de tiempo.

1- Motivation

- **Objetivo del estudio:**

Realizar la primera evaluación del desempeño de métodos de aprendizaje estadístico (AE) para el pronóstico de la inflación en Costa Rica.

- Verificar si los pronósticos de métodos de AE cumplen con las propiedades de pronósticos óptimos bajo pérdida cuadrática.
- ¿Es el desempeño de los métodos de AE superior al de sus contrapartes univariadas actualmente en uso en el BCCR?

1- Motivation

- *Aprendizaje estadístico (machine learning):*
Desarrollo y aplicación de algoritmos que permitan a máquinas mejorar su desempeño en una tarea particular al presentárseles nueva información.
- No hay marco unificado para estimación y análisis.

2- Aprendizaje estadístico y econometría

- Objetivos
 - AE: obtener algoritmos para la predicción de una variable a partir de un conjunto de otras variables. Los parámetros se establecen con respecto a una función de pérdida.
- Validación y validación cruzada
 - AE: validación cruzada fuera de muestra procura mejorar la capacidad predictiva en vez de estimar un modelo causal o estructural.
- Sobreajuste y regularización
 - AE: mayor énfasis en evitar sobreajuste. Objetivo es modelo flexible que ajuste bien, pero no a costa de capacidad predictiva.

2- Aprendizaje estadístico y econometría

- Selección de variables
 - AE: proceso es más guiado por los datos, no hay un modelo “verdadero”.
- Escalabilidad
 - AE: métodos deben poder ser aplicados con relativa facilidad cuando el tamaño del conjunto de datos se incrementa.
- Terminología
 - *Entrenamiento* → estimación.
 - “*Features*” → regresores / variables explicativas.
 - *Pesos* → coeficientes.
 - *Ejemplo o instancia* → punto de datos / observación.

3- Métodos aplicados

K vecinos más cercanos univariado

- Algoritmo de clasificación y regresión.
- Busca un conjunto de períodos similar a la historia más reciente de los datos.
- Pronóstico se basa en la evolución subsecuente de la variable de interés.

3- Métodos aplicados

K vecinos más cercanos con variables explicativas

- Principio es el mismo que con KNN univariado.
- Diferencia es que considera rezagos de otras variables además de los de Y.

3- Métodos aplicados

Bosques aleatorios

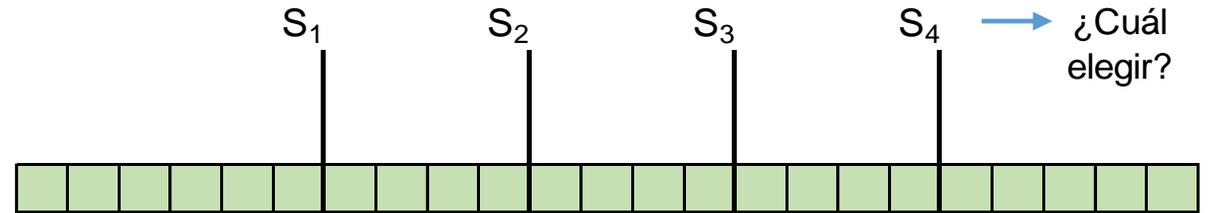
- Método no paramétrico de clasificación y regresión.
- Combinación de árboles predictivos (CART, *classification and regression trees*)

3- Métodos aplicados

Bosques aleatorios

Árbol de decisión

Valores de X_j :



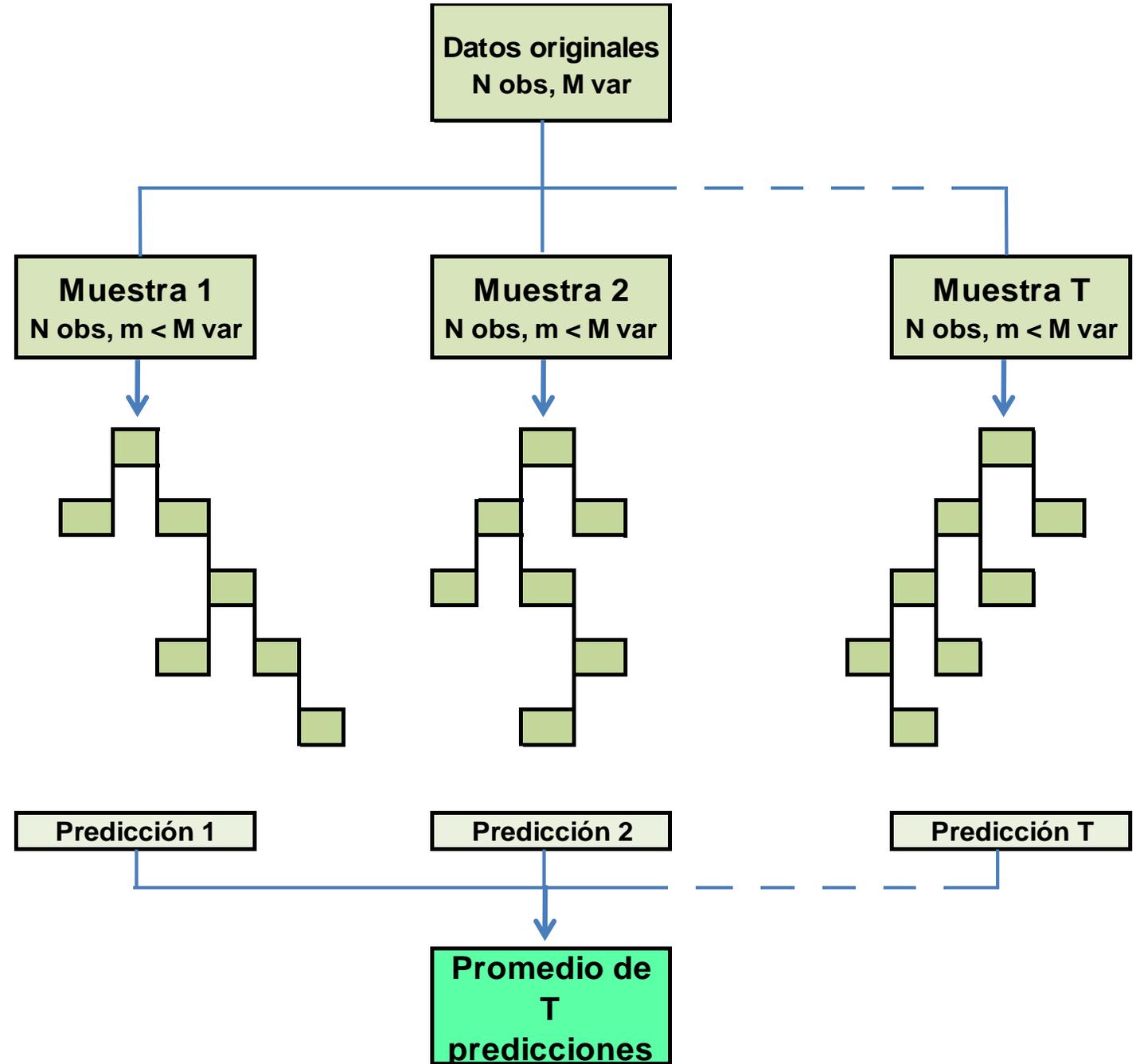
3- Métodos aplicados

Bosques aleatorios

T muestras de bootstrap
Cada una usa $m < M$
variables escogidas
al azar

T árboles de
regresión de las T
muestras de
bootstrap

T predicciones de
los T árboles



3- Métodos aplicados

Boosting

- Objetivo es mejorar la capacidad de pronóstico de métodos simples.
- Métodos de boosting no estiman un único modelo:
 - Se empieza con un modelo lineal al cual se le agrega iterativamente el regresor con la mayor contribución al ajuste global de acuerdo con el desempeño dentro de muestra.
 - No se hace ajustes a los coeficiente ya existentes.

3- Métodos aplicados

Boosting

- Ajuste repetido de mínimos cuadrados sobre los residuos.
- *Boosting del gradiente* para el modelo vigente m :
 - Obtener nuevo modelo para residuos de m : $\hat{f}^{res_m}(x)$
 - Actualizar la información vigente: $\hat{f}(x) = \hat{f}^m(x) + \lambda \hat{f}^{res_m}(x)$
 - Actualizar los residuos: $res_i = res^m - \lambda \hat{f}^{res_m}(x_i)$
 - Modelo *boosted*:
$$\hat{f}(x) = \sum_{m=1}^B \lambda \hat{f}^{res_m}(x_i)$$

3- Métodos aplicados

Long Short-Term Memory Models (LSTM)

- Mejoran considerablemente el desempeño sobre modelos de redes neuronales tradicionales:
 - Son redes *recurrentes*: permiten retroalimentación entre resultados y decisiones anteriores, y el resultado actual.
 - Permiten incorporar interdependencias de largo plazo.
 - Tienen “memoria”: un estado interno que recibe nueva información, retiene la que es útil para el pronóstico y “olvida” la que no lo es.
 - Estimación de los coeficientes en las funciones que definen cómo fluye, se retiene o se olvida la información se basa en función de pérdida (ej. error cuadrático medio).

4- Pronósticos y criterios de evaluación

Datos

- Variable a pronosticar:

tasa de inflación interanual medida mediante el IPC.

- Datos mensuales: enero-2003 a febrero-2019
- 12 rezagos de cada variable
- Variables indicadoras estacionales
- 258 variables en total

4- Pronósticos y criterios de evaluación

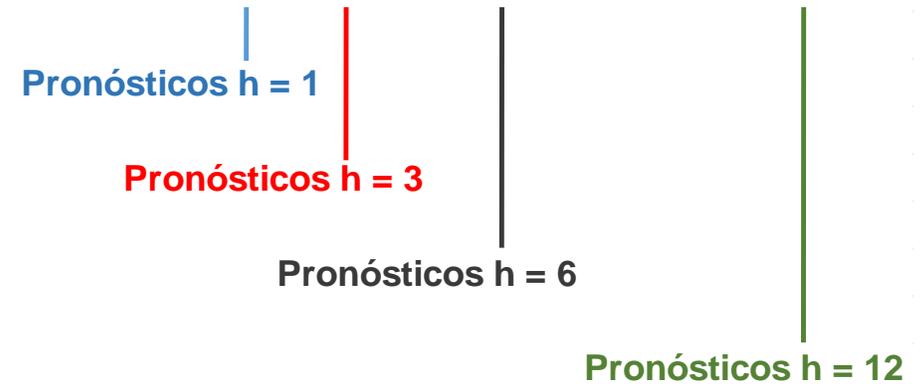
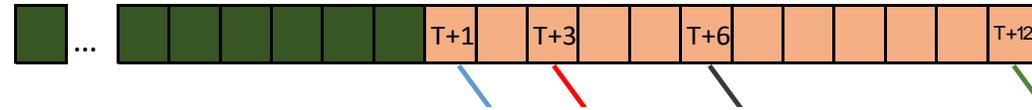
Datos

| Nombre | Descripción | Fuente |
|------------|---|---|
| IPC | Variación interanual del Índice de Precios al Consumidor (IPC), base junio de 2015 | Instituto Nacional de Estadística y Censos (INEC) |
| EXPINF12 | Expectativas de variación del IPC a 12 meses | Banco Central de Costa Rica (BCCR) |
| TCN | Tipo de cambio nominal | BCCR |
| TCR_M | Índice de Tipo de Cambio Real multilateral | BCCR |
| IMAETC | Variación interanual del Índice Mensual de Actividad Económica (IMAE), tendencia-ciclo | BCCR |
| ICFNIV | Índice de Condiciones Financieras | BCCR |
| CREDPRIVSF | Variación interanual del crédito total del sistema financiero al sector privado, moneda nacional. | BCCR |
| BASEM | Variación interanual de la base monetaria | BCCR |
| M1 | Variación interanual del M1 | BCCR |
| TPM | Tasa de política monetaria | BCCR |
| TBP | Tasa básica pasiva | BCCR |
| PRIMERATE | Tasa de referencia prime rate | BCCR |
| PETRO | Variación interanual del precio del barril de petróleo, promedio. | Pink Sheet, Banco Mundial |
| GRANOS | Variación interanual del índice de granos | Pink Sheet, Banco Mundial |
| INFSOC | Inflación de socios comerciales (variación interanual) | BCCR |
| ISMNNIV | Variación interanual del Índice de Salarios Mínimos Nominales | BCCR |
| ISMRNIV | Variación interanual del Índice de Salarios Mínimos Reales | BCCR |
| RESPIB | Resultado financiero del Gobierno Central como porcentaje del PIB | BCCR, con datos del Ministerio de Hacienda |
| DEUDAPIB | Deuda interna total como porcentaje del PIB | BCCR, con datos del Ministerio de Hacienda |

4- Pronósticos y criterios de evaluación

Procedimiento de pronóstico

Primer ejercicio



4- Pronósticos y criterios de evaluación

Criterios de evaluación

- Diebold y Lopez (1996): pronósticos óptimos bajo pérdida cuadrática:

Insesgamiento

Correlación de errores

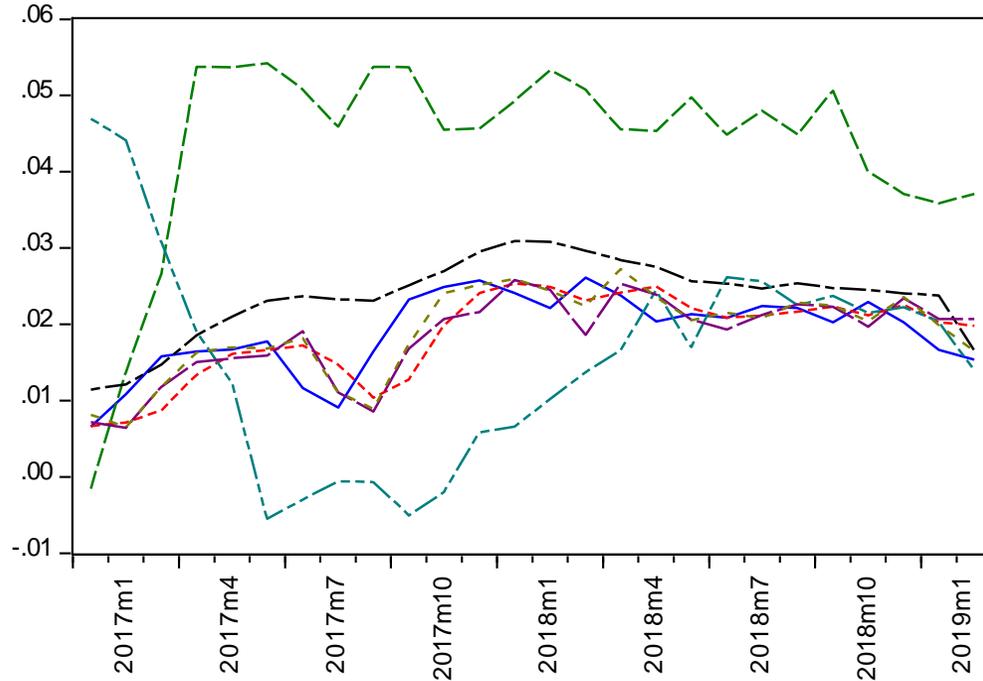
Precisión

Comportamiento de la variancia de los errores

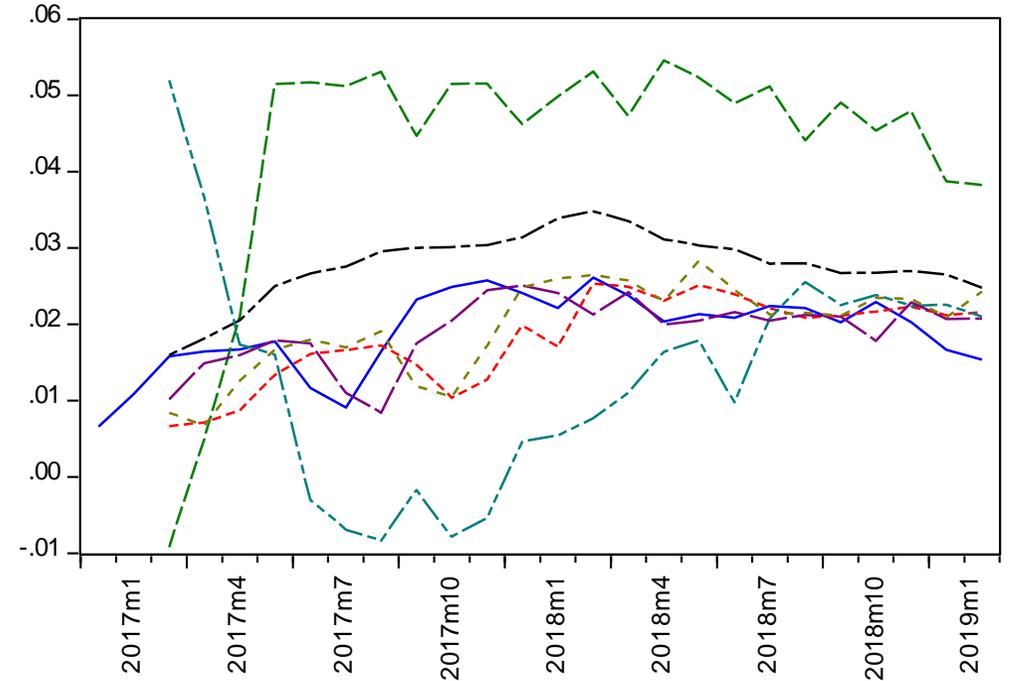
Inclusión de pronósticos

5- Resultados de la evaluación

h = 1

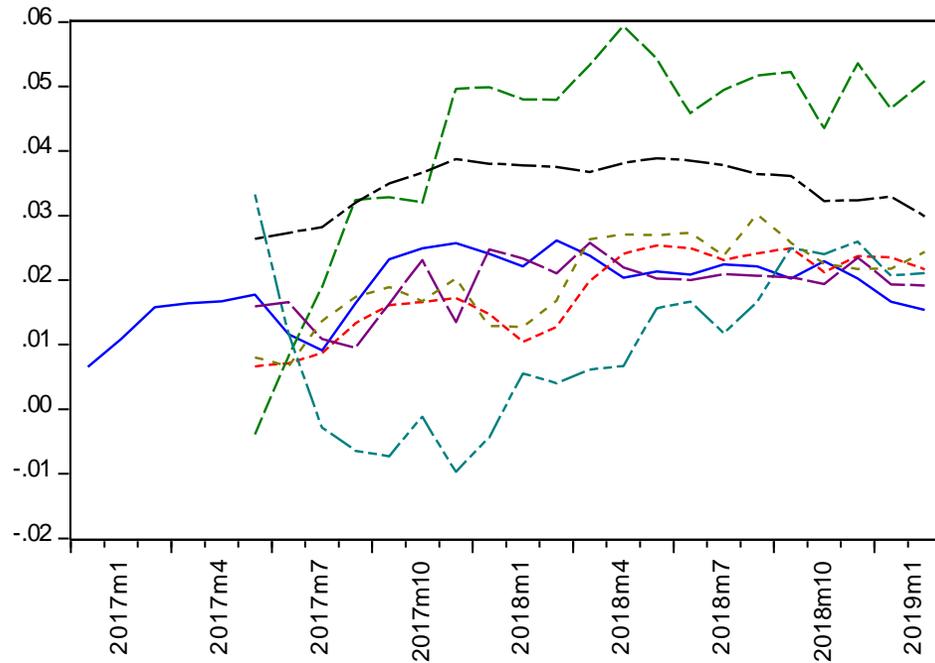


h = 3

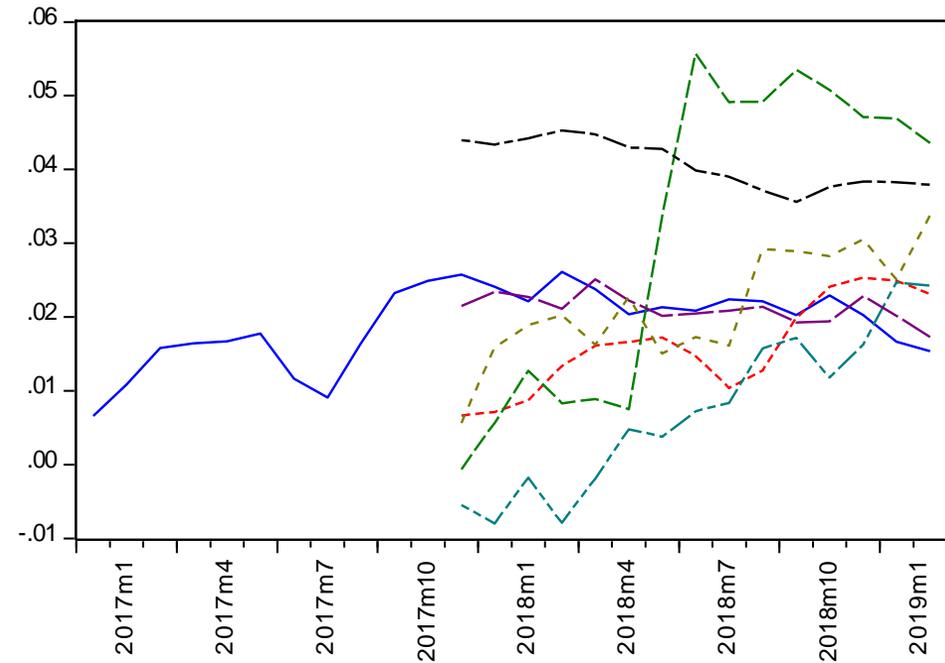


5- Resultados de la evaluación

h = 6



h = 12



5- Resultados de la evaluación

Prueba de incesgamiento de Mincer y Zarnowitz (1969)

Valor p para prueba conjunta de Wald

| | h = 1 | h = 3 | h = 6 | h = 12 |
|---------------------------------|---------------|---------------|---------------|---------------|
| KNN univariado | 0.0299 | 0.0000 | 0.0000 | 0.0000 |
| KNN con variables explicativas | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| Bosques aleatorios | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| Extreme gradient boosting | 0.0097 | 0.0000 | 0.0000 | 0.0000 |
| Modelo LSTM | 0.0454 | 0.0554 | 0.0503 | 0.7245 |
| Promedio de métodos univariados | 0.0631 | 0.0000 | 0.0000 | 0.0000 |

Valores mayores a 0.05 indican incesgamiento.

5- Resultados de la evaluación

Pruebas de autocorrelación

Valores p

| | h = 1 | h = 3 | h = 6 | h = 12 | |
|---------------------------------|---------------------|-------------------------|---------------|---------------|---------------|
| | Wilcoxon/ Dufour | Cumby & Huizinga (1992) | | | |
| KNN univariado | 0.1618 | 0.0740 | 0.3841 | 0.9331 | 0.3173 |
| KNN con variables explicativas | 0.0000 | 0.0412 | 0.4250 | 0.7268 | 0.3173 |
| Bosques aleatorios | 0.0000 | 0.0977 | 0.1797 | 0.4401 | 0.4167 |
| Extreme gradient boosting | 0.3746 | 0.1433 | 0.4072 | 0.1934 | 0.3173 |
| Modelo LSTM | 0.3603 | 0.2642 | 0.1524 | 0.5528 | 0.3173 |
| Promedio de métodos univariados | 1.0000 | 0.4717 | 0.1933 | 0.5658 | 0.3173 |

Valores mayores a 0.05 indican patrón de correlación acorde con pronósticos óptimos bajo pérdida cuadrática

5- Resultados de la evaluación

Precisión *RECM y Theil*

| | RECM | | | | Theil ^{1/} | | | |
|---------------------------------|--------|--------|--------|--------|---------------------|--------|--------|---------|
| | h = 1 | h = 3 | h = 6 | h = 12 | h = 1 | h = 3 | h = 6 | h = 12 |
| KNN univariado | 0.0039 | 0.0062 | 0.0067 | 0.0100 | 0.1006 | 0.1587 | 0.1676 | 0.25743 |
| KNN con variables explicativas | 0.0267 | 0.0277 | 0.0256 | 0.0244 | 0.4111 | 0.4174 | 0.3910 | 0.41104 |
| Bosques aleatorios | 0.0057 | 0.0089 | 0.0145 | 0.0193 | 0.1295 | 0.1834 | 0.2612 | 0.30792 |
| Extreme gradient boosting | 0.0044 | 0.0072 | 0.0072 | 0.0144 | 0.1130 | 0.1851 | 0.1819 | 0.38501 |
| Modelo LSTM | 0.0039 | 0.0036 | 0.0042 | 0.0024 | 0.1003 | 0.0892 | 0.1027 | 0.05648 |
| Promedio de métodos univariados | 0.0032 | 0.0061 | 0.0065 | 0.0094 | 0.0805 | 0.1487 | 0.1552 | 0.21178 |

Valores más bajos indican mejor ajuste

5- Resultados de la evaluación

Valores p de prueba de Harvey, Leybourne y Newbold (1997)

h = 1

| | KNN univariado | KNN con variables explicativas | Bosques aleatorios | Extreme gradient boosting | Modelo LSTM | Promedio de métodos univariados |
|---|----------------|--------------------------------|--------------------|---------------------------|---------------|---------------------------------|
| KNN univariado | | 1.0000 | 0.9252 | 0.6721 | 0.4751 | 0.0743 |
| KNN con variables explicativas | 0.0000 | | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| Bosques aleatorios | 0.0748 | 1.0000 | | 0.1709 | 0.0484 | 0.0172 |
| Extreme gradient boosting | 0.3279 | 1.0000 | 0.8291 | | 0.3043 | 0.0998 |
| Modelo LSTM | 0.5249 | 1.0000 | 0.9516 | 0.6957 | | 0.0057 |
| Promedio de métodos univariados | 0.9257 | 1.0000 | 0.9828 | 0.9002 | 0.9943 | |
| Casos en que error de columna es menor a error de fila: | 1 | 0 | 1 | 1 | 2 | 3 |

h = 3

| | KNN univariado | KNN con variables explicativas | Bosques aleatorios | Extreme gradient boosting | Modelo LSTM | Promedio de métodos univariados |
|---|----------------|--------------------------------|--------------------|---------------------------|---------------|---------------------------------|
| KNN univariado | | 1.0000 | 0.8898 | 0.7423 | 0.0579 | 0.3927 |
| KNN con variables explicativas | 0.0000 | | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| Bosques aleatorios | 0.1102 | 1.0000 | | 0.1781 | 0.0031 | 0.0762 |
| Extreme gradient boosting | 0.2577 | 1.0000 | 0.8219 | | 0.0055 | 0.2259 |
| Modelo LSTM | 0.9421 | 1.0000 | 0.9969 | 0.9945 | | 0.9645 |
| Promedio de métodos univariados | 0.6073 | 1.0000 | 0.9238 | 0.7741 | 0.0355 | |
| Casos en que error de columna es menor a error de fila: | 1 | 0 | 1 | 1 | 4 | 1 |

Valores menores a 0.05 indican que pronóstico de columna es más preciso que el de la fila

5- Resultados de la evaluación

Valores p de prueba de Harvey, Leybourne y Newbold (1997)

h = 6

| | KNN univariado | KNN con variables explicativas | Bosques aleatorios | Extreme gradient boosting | Modelo LSTM | Promedio de métodos univariados |
|---|----------------|--------------------------------|--------------------|---------------------------|---------------|---------------------------------|
| KNN univariado | | 0.9995 | 1.0000 | 0.5882 | 0.0535 | 0.3980 |
| KNN con variables explicativas | 0.0005 | | 0.0041 | 0.0007 | 0.0004 | 0.0004 |
| Bosques aleatorios | 0.0000 | 0.9959 | | 0.0000 | 0.0000 | 0.0000 |
| Extreme gradient boosting | 0.4118 | 0.9993 | 1.0000 | | 0.0506 | 0.3657 |
| Modelo LSTM | 0.9465 | 0.9996 | 1.0000 | 0.9494 | | 0.9721 |
| Promedio de métodos univariados | 0.6020 | 0.9996 | 1.0000 | 0.6343 | 0.0279 | |
| Casos en que error de columna es menor a error de fila: | 2 | 0 | 1 | 2 | 3 | 2 |

h = 12

| | KNN univariado | KNN con variables explicativas | Bosques aleatorios | Extreme gradient boosting | Modelo LSTM | Promedio de métodos univariados |
|---|----------------|--------------------------------|--------------------|---------------------------|---------------|---------------------------------|
| KNN univariado | | 0.9903 | 0.9995 | 0.7712 | 0.0420 | 0.4078 |
| KNN con variables explicativas | 0.0097 | | 0.1374 | 0.0046 | 0.0013 | 0.0031 |
| Bosques aleatorios | 0.0005 | 0.8626 | | 0.1543 | 0.0000 | 0.0001 |
| Extreme gradient boosting | 0.2288 | 0.9954 | 0.8457 | | 0.0495 | 0.1800 |
| Modelo LSTM | 0.9580 | 0.9987 | 1.0000 | 0.9505 | | 0.9690 |
| Promedio de métodos univariados | 0.5923 | 0.9969 | 0.9999 | 0.8200 | 0.0310 | |
| Casos en que error de columna es menor a error de fila: | 2 | 0 | 0 | 1 | 5 | 2 ²⁶ |

Valores menores a 0.05 indican que pronóstico de columna es más preciso que el de la fila

5- Resultados de la evaluación

Predicción de la dirección de cambios en la inflación
Resaltado si es mayor a 50%

| | h = 1 | h = 3 | h = 6 | h = 12 |
|---------------------------------|---------------|---------------|---------------|---------------|
| KNN univariado | 0.5600 | 0.6087 | 0.5500 | 0.4286 |
| KNN con variables explicativas | 0.6400 | 0.6522 | 0.3500 | 0.3571 |
| Bosques aleatorios | 0.6800 | 0.6522 | 0.6000 | 0.7143 |
| Extreme gradient boosting | 0.4800 | 0.6522 | 0.5500 | 0.4286 |
| Modelo LSTM | 0.6000 | 0.5652 | 0.5500 | 0.5000 |
| Promedio de métodos univariados | 0.4800 | 0.6087 | 0.4500 | 0.2857 |

5- Resultados de la evaluación

- Pruebas de Chong and Hendry:
 - A horizontes más cortos, los modelos univariados incluyen a KNN y boosting.
 - A horizontes mayores los pronósticos el modelo LSTM muestran mejor desempeño: únicos que incluyen otros pronósticos, particularmente para $h=3$ y $h=12$.

5- Resultados de la evaluación

Variances of forecast errors

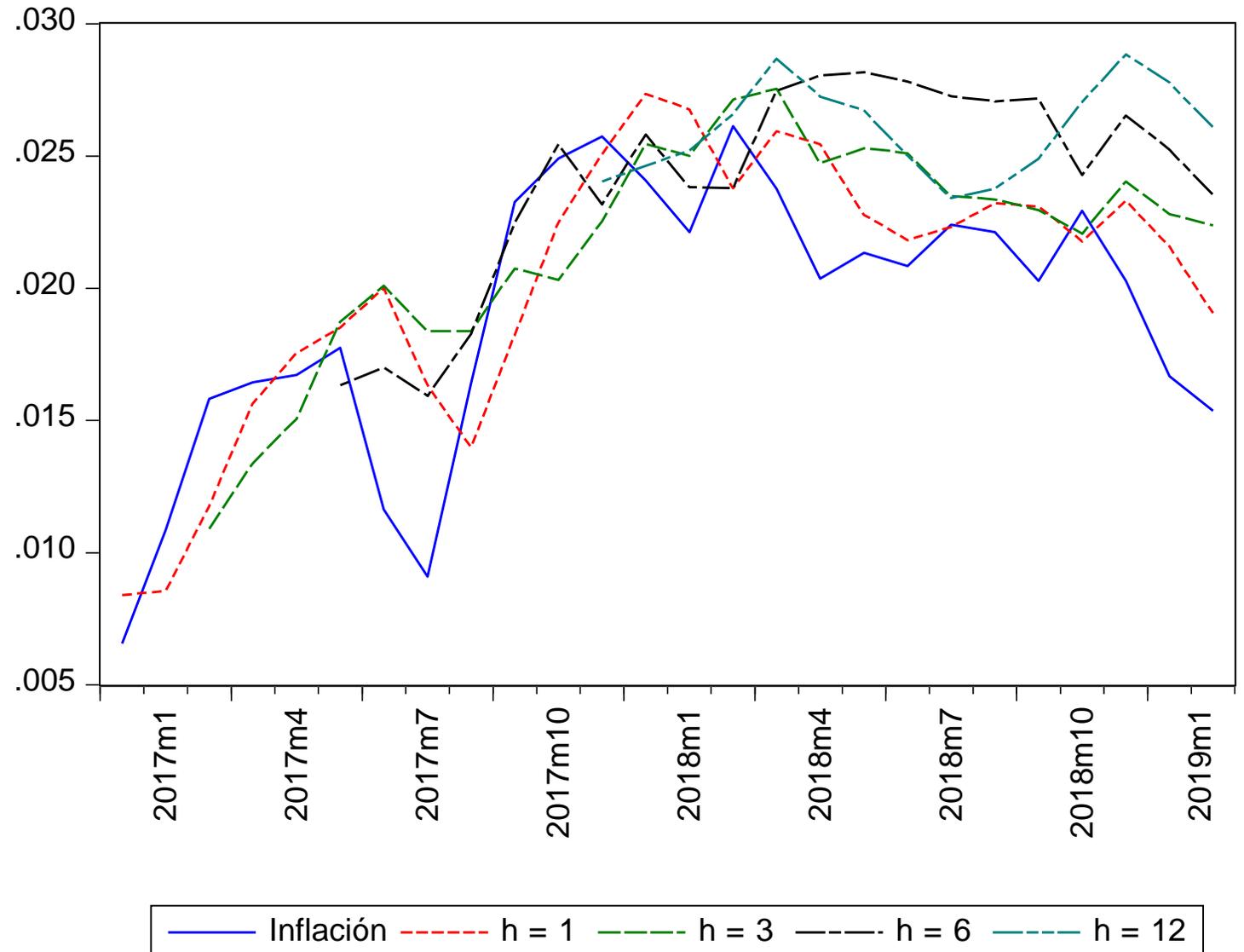
| | h = 1 | h = 3 | h = 6 | h = 12 |
|---------------------------------|--------------|--------------|--------------|---------------|
| KNN univariado | 0.00002 | 0.00004 | 0.00004 | 0.00007 |
| KNN con variables explicativas | 0.00012 | 0.00023 | 0.00021 | 0.00054 |
| Bosques aleatorios | 0.00001 | 0.00002 | 0.00001 | 0.00001 |
| Extreme gradient boosting | 0.00002 | 0.00005 | 0.00005 | 0.00011 |
| Modelo LSTM | 0.00002 | 0.00001 | 0.00002 | 0.00001 |
| Promedio de métodos univariados | 0.00001 | 0.00004 | 0.00004 | 0.00010 |

6- Combinación de pronósticos

Capistrán y

Timmermann (2009):

- LSTM
- KNN univariados
- Bosques aleatorios



6- Combinación de pronósticos

Evaluación de combinación de pronósticos *Varias pruebas*

| | Valor p para hipótesis nula de inestabilidad | Predicción de dirección de cambios en inflación (% éxito) | Valores p para pruebas de autocorrelación ^{1/} | Variaciones de errores de pronóstico |
|-------------|--|---|---|--------------------------------------|
| h=1 | 1.00 | 0.6400 | 0.2940 | 0.000010 |
| h=3 | 1.00 | 0.6957 | 0.2832 | 0.000013 |
| h=6 | 1.00 | 0.5000 | 0.5932 | 0.000012 |
| h=12 | 1.00 | 0.5714 | 0.3173 | 0.000008 |

- Inestable
- Buena predicción de dirección de cambios
- Errores de pronóstico con patrones de correlación apropiados

6- Combinación de pronósticos

P-values for the HLN (1997) test

Resaltado si el error de pronóstico es menor que el de la fila

| | Combinación | | | |
|--------------------------------|---------------|---------------|---------------|---------------|
| | h = 1 | h = 3 | h = 6 | h = 12 |
| KNN univariado | 0.0847 | 0.0490 | 0.0234 | 0.0457 |
| KNN con variables explicativas | 0.0000 | 0.0000 | 0.0003 | 0.0013 |
| Bosques aleatorios | 0.0091 | 0.0015 | 0.0000 | 0.0000 |
| Extreme gradient boosting | 0.0999 | 0.0037 | 0.0226 | 0.0526 |
| Modelo LSTM | 0.0112 | 0.4333 | 0.1966 | 0.6729 |
| Promedio 1-5 | 0.4132 | 0.0327 | 0.0029 | 0.0288 |

7- Conclusiones

- El objetivo de este estudio fue realizar una primera evaluación de la capacidad de varios métodos de aprendizaje automático para pronosticar la inflación de Costa Rica.
- El modelo LSTM genera los pronósticos con el mejor desempeño: insesgados para $h > 1$, más precisos, abarcan a la mayoría de otros pronósticos.
- Siguiendo en desempeño: KNN univariados, y en menor medida bosques aleatorios y boosting extremo del gradiente.
- Combinación de pronósticos mejora el desempeño sobre pronósticos individuales a todos los horizontes, y es mejor que los métodos univariados.
 - Insesgada
 - Errores con patrones de correlación adecuados
 - Mejora la capacidad de pronóstico a todos los horizontes (nivel y dirección)

7- Conclusiones

- Resultados promisorios de la implementación de métodos de aprendizaje automático para el pronóstico en BCCR.
- Trabajo siguiente:
 - Mejorar la aplicación de métodos con desempeño insatisfactorio en el estudio.
 - Potencial extensión del trabajo para incluir métodos adicionales.
 - Pronóstico en tiempo real del crecimiento del producto.

JORNADAS DE INVESTIGACIÓN ECONÓMICA

Departamento de Investigación Económica

Diciembre, 2019

3- Métodos aplicados

Bosques aleatorios

Árbol de decisión

- Para cada variable j y valor s , se definen los siguientes conjuntos:

R1: valores de X menores a s

R2: valores de X mayores o iguales a s

- Se encuentra los valores de j y s que minimizan suma de desviaciones al cuadrado de cada valor de X con respecto al promedio del grupo al que pertenece.

$$\sum_{i:x_i \in R1(j,s)} (y_i - \bar{y}_{R1})^2 + \sum_{i:x_i \in R2(j,s)} (y_i - \bar{y}_{R2})^2$$

- Tras la primera división, se utiliza el mismo criterio para volver a dividir $R1$ y $R2$.
- Se detiene proceso cuando quedan p observaciones en cada región (tamaño terminal de nodo)

3- Métodos aplicados

Bosques aleatorios

- Muy probable que haya sobreajuste.
- Alta variancia si los árboles en cada división son muy distintos.
- *Bagging*:
Promedio de los resultados de los árboles para muchas muestras de bootstrapping. Se calcula una medida de error con base en las observaciones no incluidas en la muestra de bootstrapp (*out-of-bag error*).
- Interpretación complicada a través de los árboles
 - Importancia de cada regresor: calcular el promedio para todos los árboles del monto en que la SRC cae debido a divisions en ese regresor.
 - Valor alto → regresor importante

3- Métodos aplicados

Bosques aleatorios

- Muy probable que haya sobreajuste.
- Interpretación complicada a través de los árboles
 - Importancia de cada regresor: calcular el promedio para todos los árboles del monto en que la SRC cae debido a divisiones en ese regresor.
 - Valor alto → regresor importante

3- Métodos aplicados

Modelos Long Short-Term Memory (LSTM)

- Hochreiter y Schmidhuber (1997), mejoran desempeño sobre modelos de redes neuronales tradicionales.

- **Estados:**

- *Datos de entrada:* X

- *Estado oculto:* llamado h , valores de la capa escondida previa

- *Estado de insumo:* combinación lineal de estado oculto y de datos de insumo en período actual.

$$i_t = \sigma(W_{ix}X_t + W_{ih}h_{t-1})$$

W_{ix} : matriz de ponderaciones W_{ih} : matriz de transición entre estados

- *Estado interno:* denotado m , valores que actúan como la memoria del modelo.

3- Métodos aplicados

Modelos Long Short-Term Memory (LSTM)

- **Puertas:** valores usados para decidir el flujo de Información entre estados.
 - *Puerta de insumos:* “decide” si estado de insumo entra en el estado interno.

$$g_t = \sigma(W_{gi}i_t)$$

- *Puerta de olvido:* decide si estado interno “olvida” el estado interno anterior

$$f_t = \sigma(W_{fi}i_t)$$

- *Puerta de resultado:* decide si el estado interno pasa su valor a resultado y al estado interno del siguiente período.

$$o_t = \sigma(W_{oi}i_t)$$

- **Flujo de información:**

$$m_t = g_t \odot i_t + f_t \odot m_{t-1}$$

$$h_t = o_t \odot m_t$$

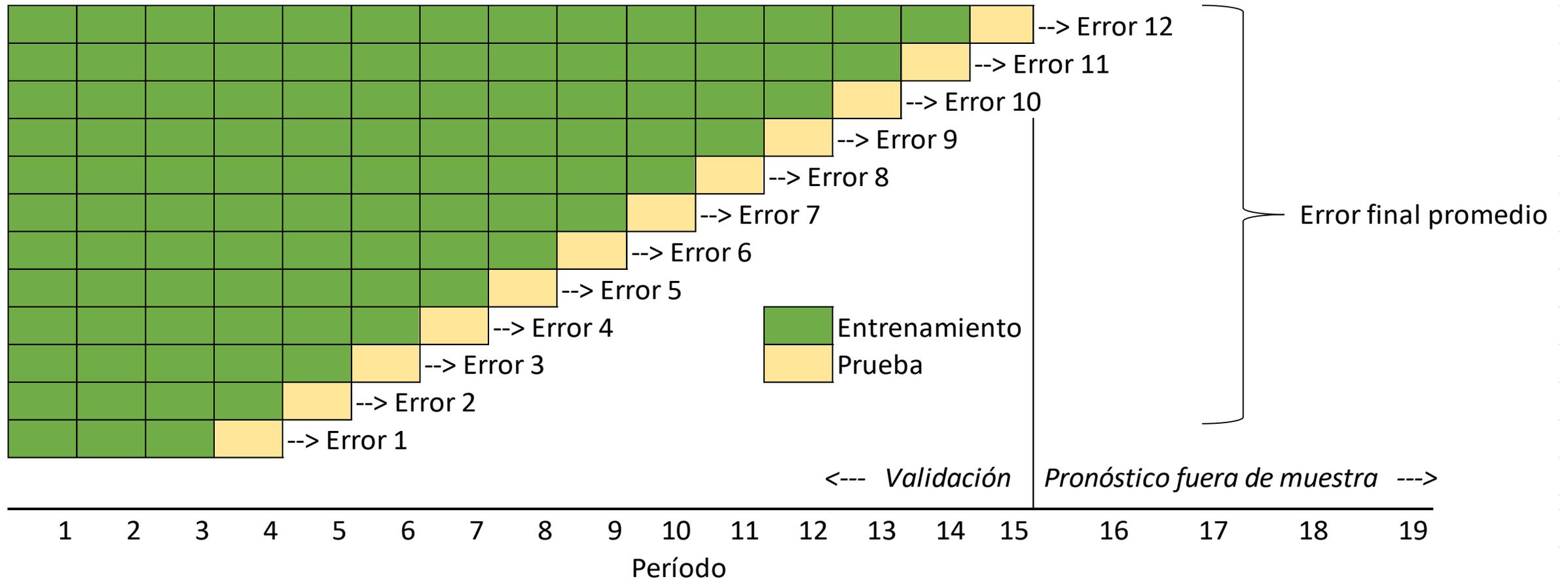
4- Pronósticos y criterios de evaluación

Procedimiento de pronóstico

- Muestras de estimación:
 - Enero 2003 a diciembre 2016 → 12 pronósticos multihorizonte
 - Se aumenta la muestra de estimación un mes a la vez.
- Calibración donde sea necesaria (validación cruzada).
- A partir de los pronósticos multihorizonte, se forman series de horizonte fijo a 1, 3, 6 y 12 meses.

4- Pronósticos y criterios de evaluación

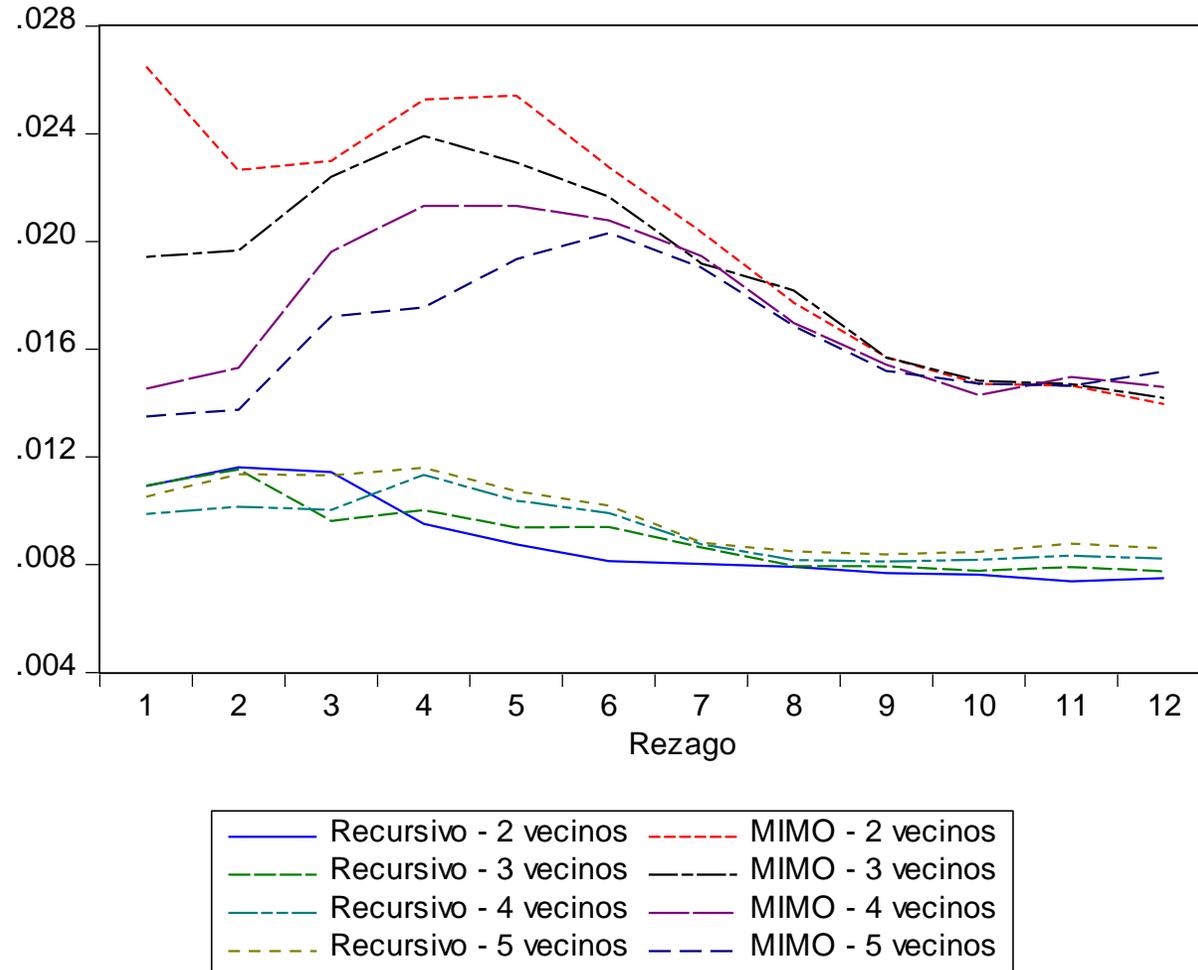
Validación cruzada



4- Pronósticos y criterios de evaluación

KNN univariado

RECM para modelos KNN univariados



4- Pronósticos y criterios de evaluación

KNN con variables explicativas

- 11 rezagos de todas las variables
- Variables incluidas:
 - Expectativas de inflación a 12 meses (t)
 - Tipo de cambio (variación interanual en t)
 - Base monetaria (variación interanual, t-4)
 - Tasa de política monetaria (t-9)
 - Precio del petróleo (variación interanual, t-8)
 - Precio de los granos (variación interanual, t-6).
- Mínima RECM con 5 vecinos
- Agregación mediante promedio ponderado por inverso de la distancia.

4- Pronósticos y criterios de evaluación

Bosques aleatorios

- *Número de árboles:* se estableció en 100 para todos los ejercicios. Ver Breiman (2001) , Probst and Boulesteix (2018)
- *Tamaño de los nodos terminales:* 5 observaciones (valor usual en la mayoría de aplicaciones).
- *Número de variables en cada árbol:* este parámetro se calibró para cada ejercicio de pronóstico, encontrando el número de variables que minimiza el error *out-of-bag*.

4- Pronósticos y criterios de evaluación

Boosting

- *Boosting extremo del gradiente*: Chen and Guestrin (2016), adaptación regularizada de boosting de gradiente.
- RECM como métrica de evaluación, selección considerando una variable a la vez.
- Parámetros a calibrar:
 - *nrounds*: max. número de iteraciones para convergencia del algoritmo del descenso del gradiente.
 - *lambda*: parámetro para evitar sobreajuste, controlando un proceso de regularización en los coeficientes, equivalente a regresión Ridge (L2).
 - *alfa*: para evitar sobreajuste, proceso de regularización (L1) similar a Lasso.

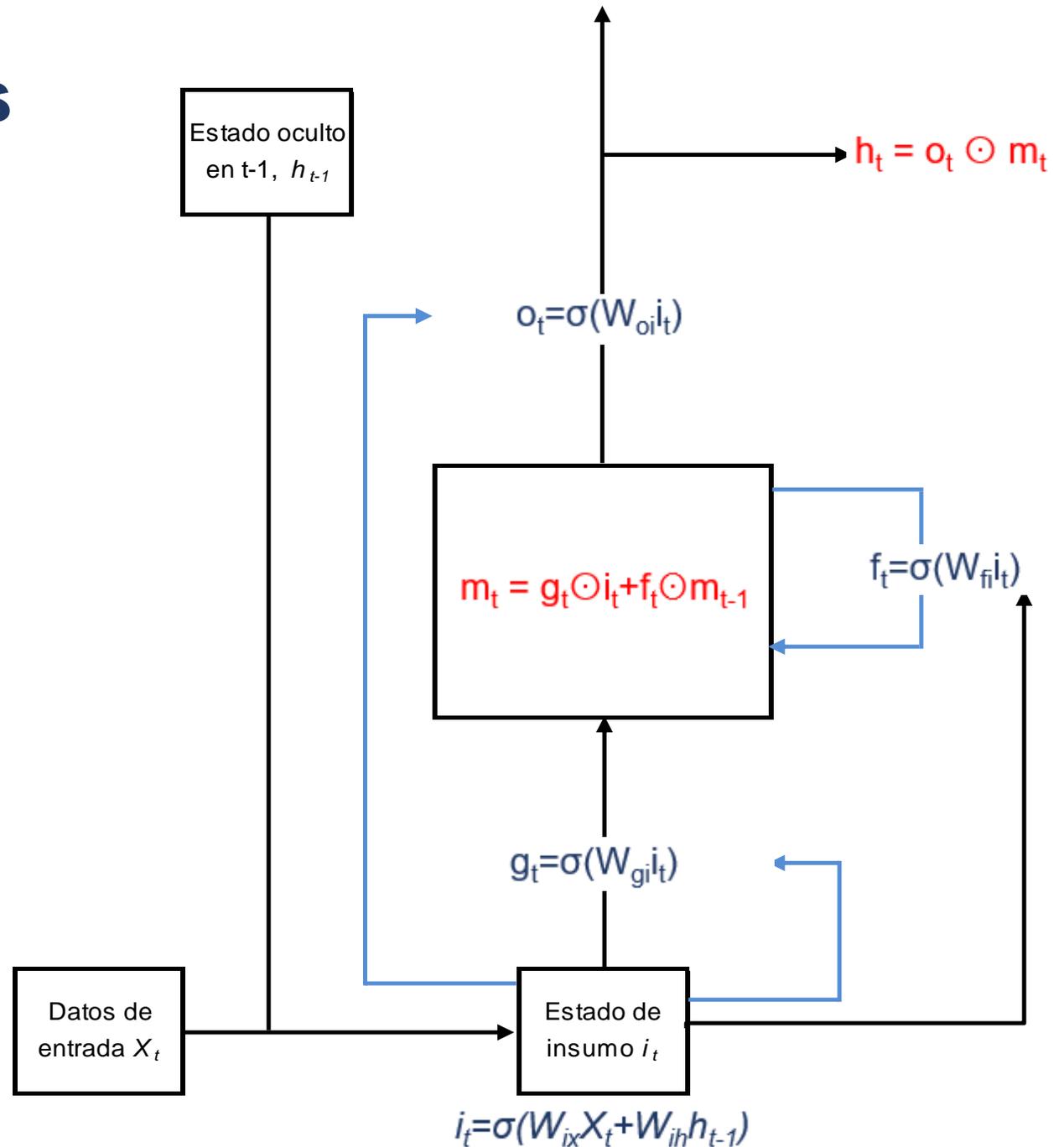
4- Pronósticos y criterios de evaluación

LSTM

- Métrica para evaluación durante entrenamiento: precisión (RECM).
- Para optimización mediante descenso del gradiente
 - Tasa de aprendizaje usual: 0.02.
 - Decaimiento usual de la tasa de aprendizaje: $1e-6$.
 - Tangente hiperbólica como función de activación.

3- Métodos aplicados

Long Short-Term Memory Models



4- Pronósticos y criterios de evaluación

Criterios de evaluación

- Diebold y Lopez (1996): pronósticos óptimos bajo pérdida cuadrática

Insesgamiento

$$y_{t+s} = \alpha + \beta \hat{y}_{t+s} + \varepsilon_{t+s}$$

Correlación de errores

- No autocorrelacionados para $h=1$
- A lo sumo siguen un MA($h-1$) para $h>1$

Precisión

- RECM & Theil
- Harvey, Leybourne y Newbold (1997)
- % de dirección de cambios

4- Pronósticos y criterios de evaluación

Criterios de evaluación

Variaciones de los errores de pronóstico

- Variancia no debería disminuir conforme aumenta horizonte.

Inclusión de pronósticos

- Chong y Hendry (1986)

$$y_{t+s} = \beta_0 + \beta_1 \hat{y}_{t+s,1} + \beta_2 \hat{y}_{t+s,2} + \varepsilon_{t+s}$$

$(\beta_0 \ \beta_1 \ \beta_2) = (0 \ 1 \ 0)$ pronóstico 1 abarca al 2

$(\beta_0 \ \beta_1 \ \beta_2) = (0 \ 0 \ 1)$ pronóstico 2 abarca al 1