

PRONÓNSTICO PARA EL TIPO DE CAMBIO UTILIZANDO REDES NEURONALES CONVOLUCIONALES

Presentado por:

David Gabriel

“No temo a las computadoras. Temo no tenerlas”

-Isaac Asimov

Antecedentes

Históricamente, el tipo de cambio ha sido muy impredecible. El modelo de caminata aleatoria es un *benchmark* difícil de superar, tanto para modelos autoregresivos, como para modelos lineales multivariados.

Objetivo

Pronosticar el tipo de cambio “*one-step ahead*” (pronosticar el tipo de cambio en $t + 1$ utilizando valores pasados del mismo tipo de cambio hasta t) con redes neuronales convolucionales, y comparar su rendimiento con modelos *benchmark*. La hipótesis nula es que los modelos con redes neuronales no son mejores* que el de caminata aleatoria.

INTRODUCCIÓN A LAS CONVOLUCIONES

The background features abstract, overlapping geometric shapes in various shades of green, ranging from light lime to dark forest green. These shapes are primarily located on the right side of the frame, creating a modern, layered effect. The text is centered in a clean, green, serif font.

¿Qué es una convolución?

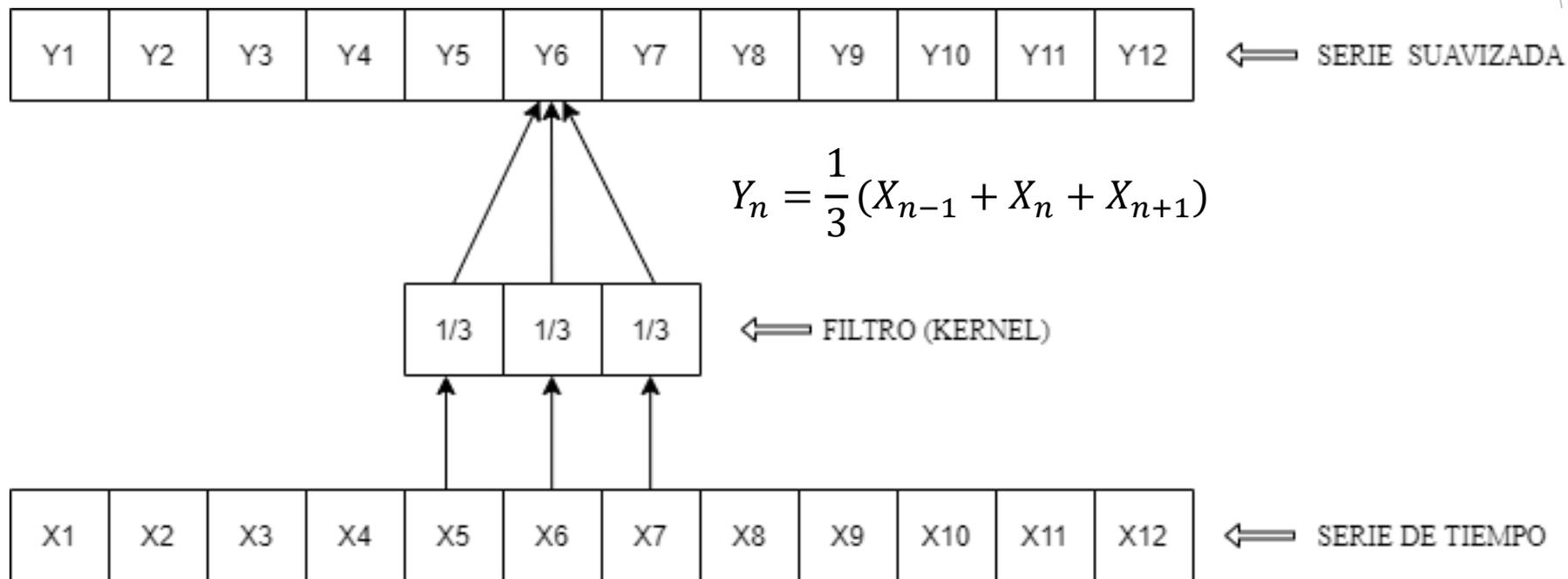
Sean f y g dos funciones discretas, se define una convolución como:

$$f \star g (t) = \sum_{\tau=-L}^L f(t - \tau)g(\tau)$$

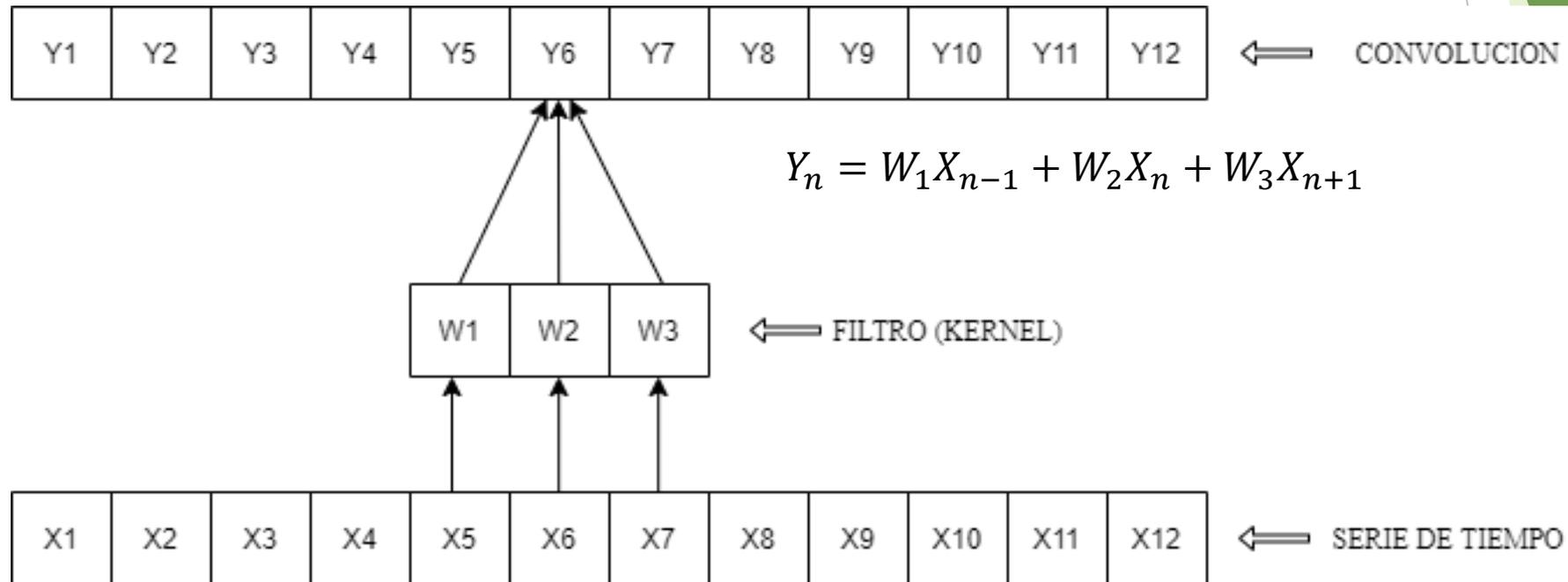
Ejemplo 1: Una media móvil es un caso específico de convolución



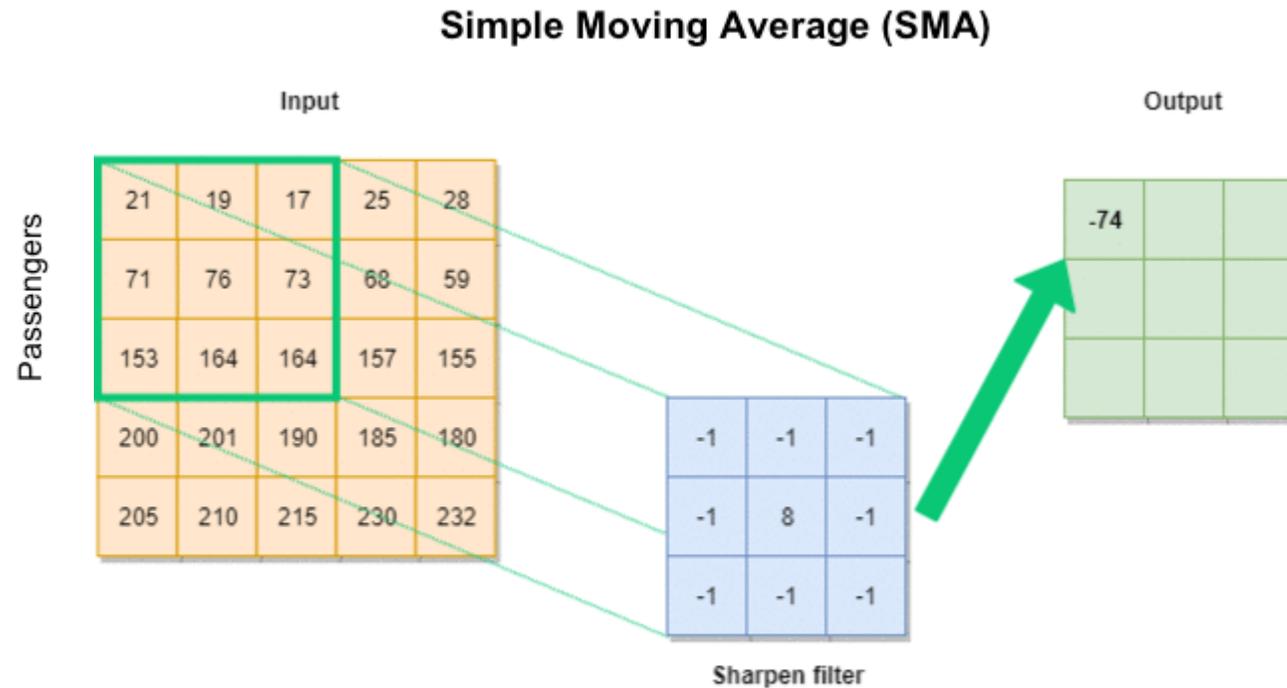
En una media móvil, los valores nuevos de una serie, son el promedio de los valores en su vecindad.



En una convolución, reemplazamos la media con cualquier combinación lineal.



Ejemplo 2: Convolución 2D



AIGeekProgrammer.com © 2019

Fuente: <https://aigeekprogrammer.com/wp-content/uploads/2019/12/CNN-filter-animation-1.gif>

¿Por qué convoluciones?

Las convoluciones son particularmente útiles para extraer patrones entre los datos.

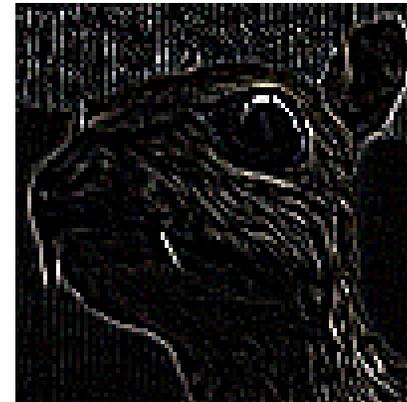
Input image



Convolution
Kernel

$$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

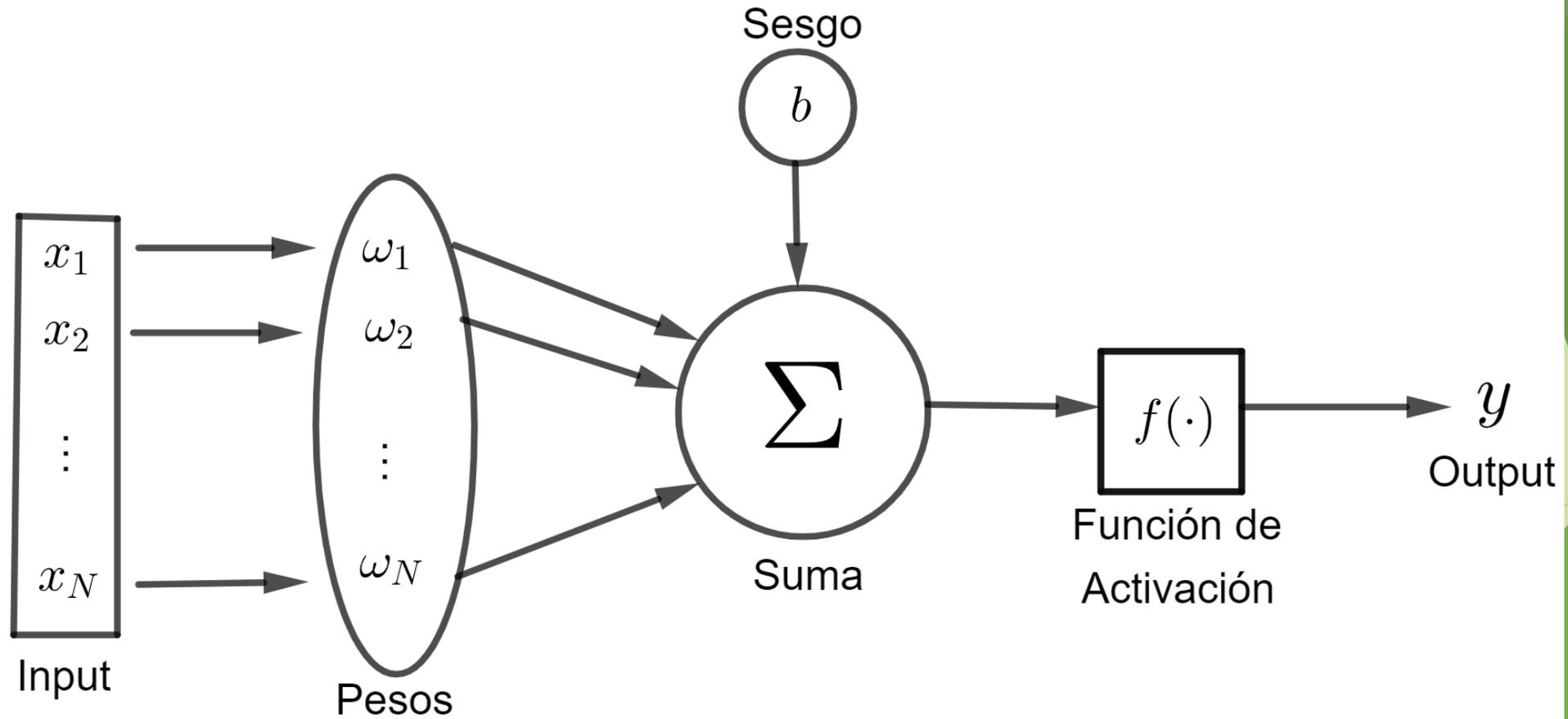
Feature map



INTRODUCCIÓN A LAS REDES NEURONALES

The background features abstract, overlapping geometric shapes in various shades of green, ranging from light lime to dark forest green. These shapes are primarily located on the right side of the frame, creating a modern, layered effect. The text is centered on a white background.

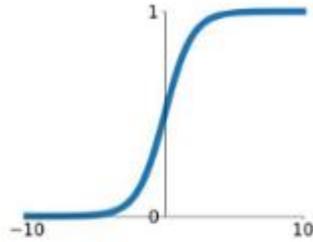
El Perceptrón



Funciones de Activación

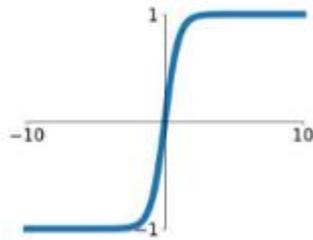
Sigmoid

$$\sigma(x) = \frac{1}{1+e^{-x}}$$



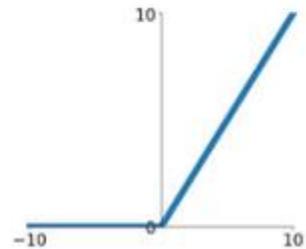
tanh

$$\tanh(x)$$



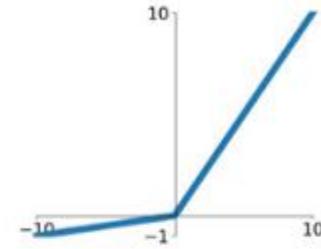
ReLU

$$\max(0, x)$$



Leaky ReLU

$$\max(0.1x, x)$$

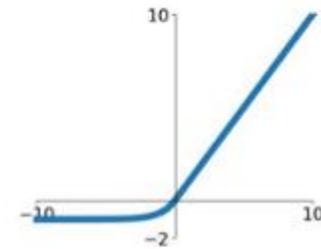


Maxout

$$\max(w_1^T x + b_1, w_2^T x + b_2)$$

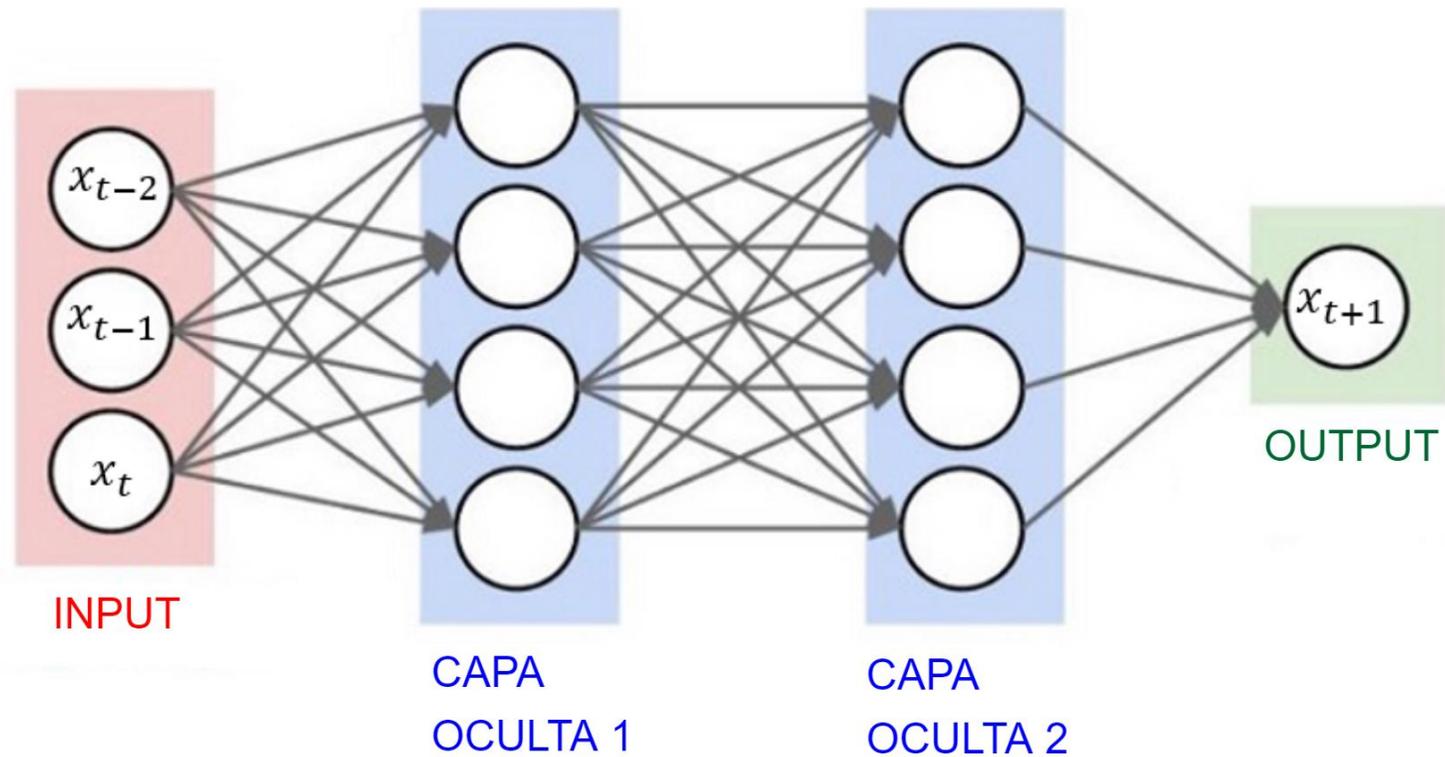
ELU

$$\begin{cases} x & x \geq 0 \\ \alpha(e^x - 1) & x < 0 \end{cases}$$

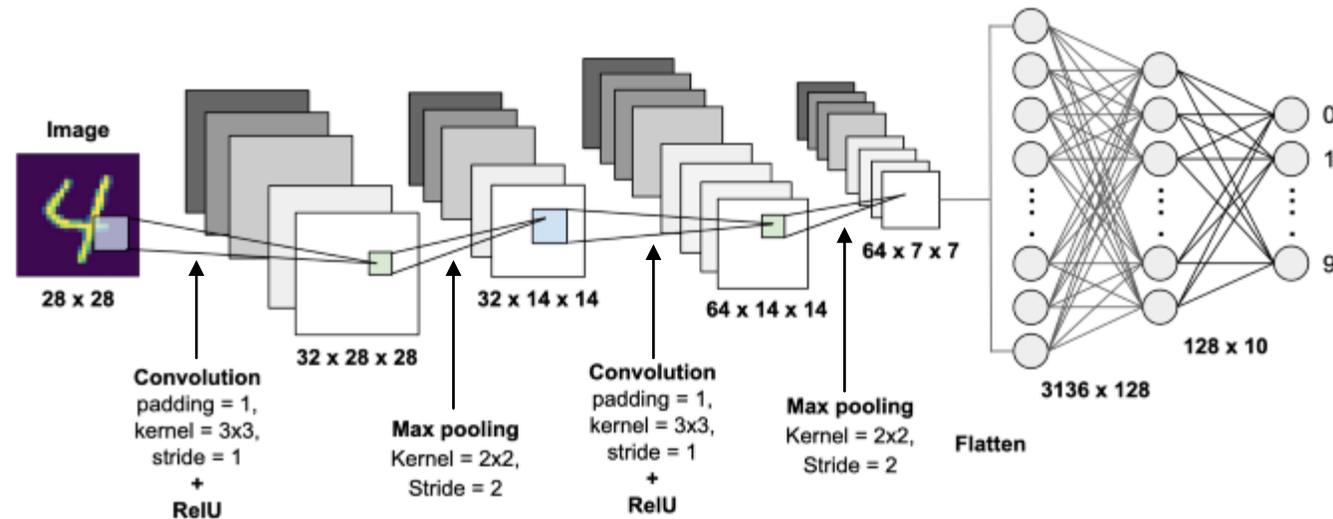


Fuente: <https://mc.ai/complete-guide-of-activation-functions/>

¿Qué es una red neuronal?

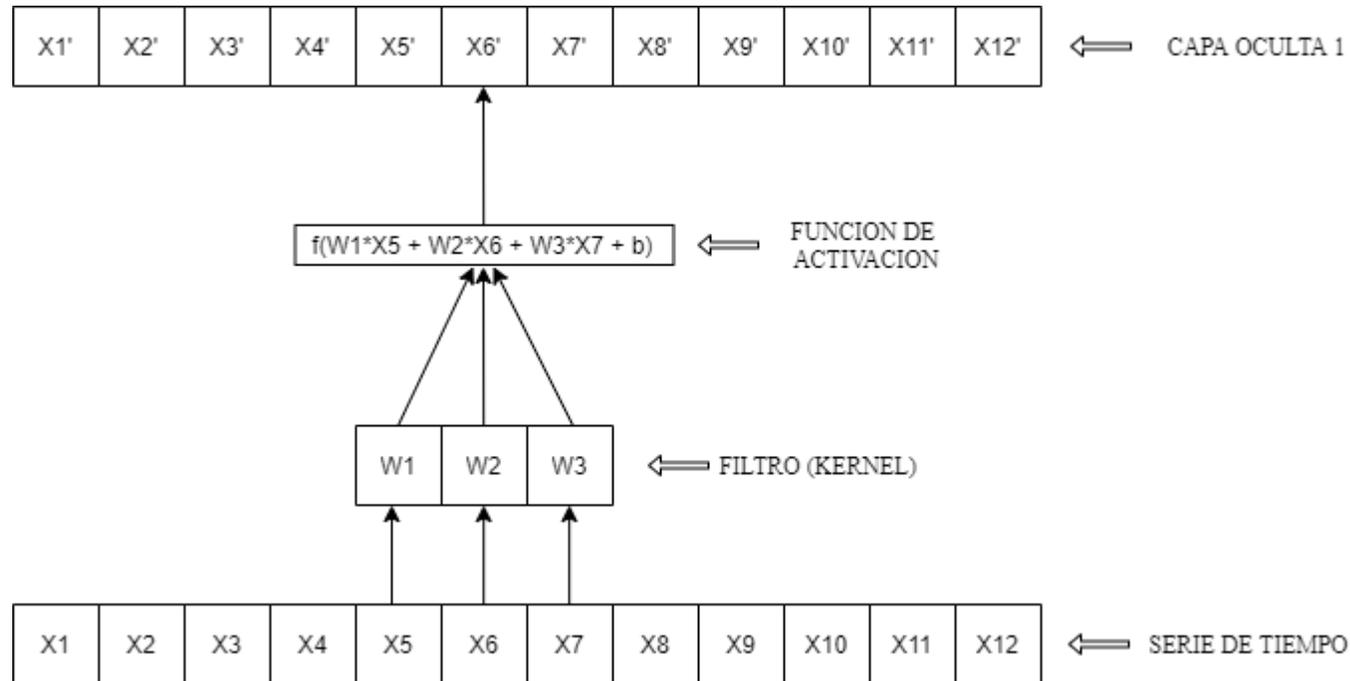


Introducción a las redes neuronales convolucionales



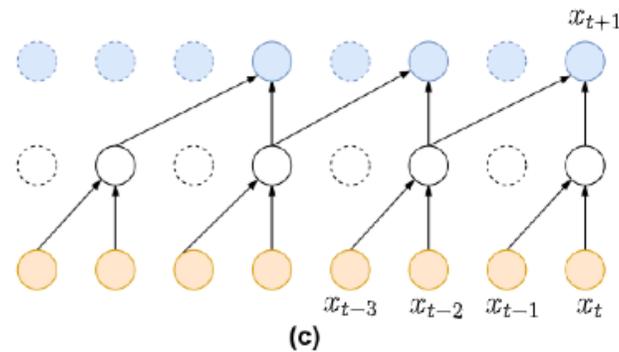
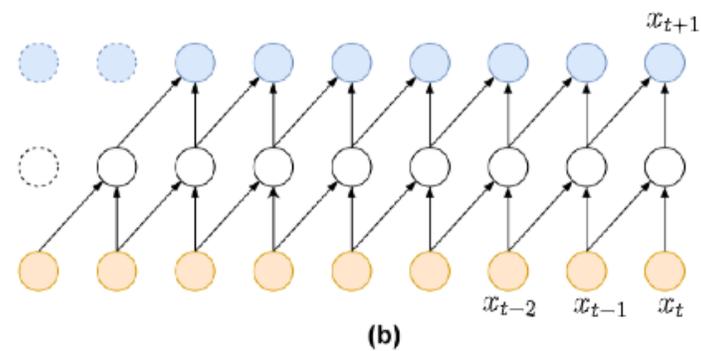
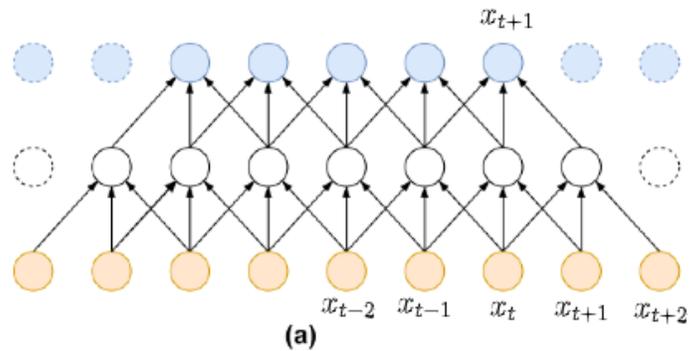
Fuente: <https://becominghuman.ai/building-a-convolutional-neural-network-cnn-model-for-image-classification-116f77a7a236>

En el caso de una serie de tiempo

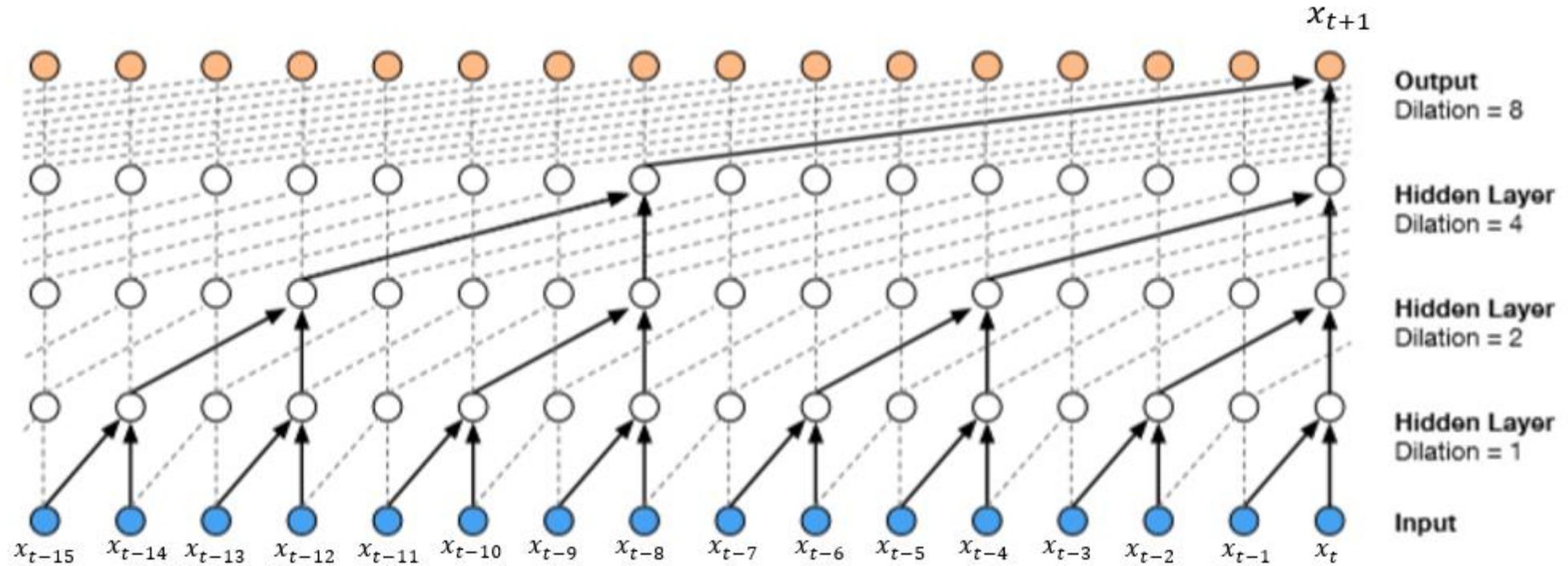


Problemas con las redes ordinarias

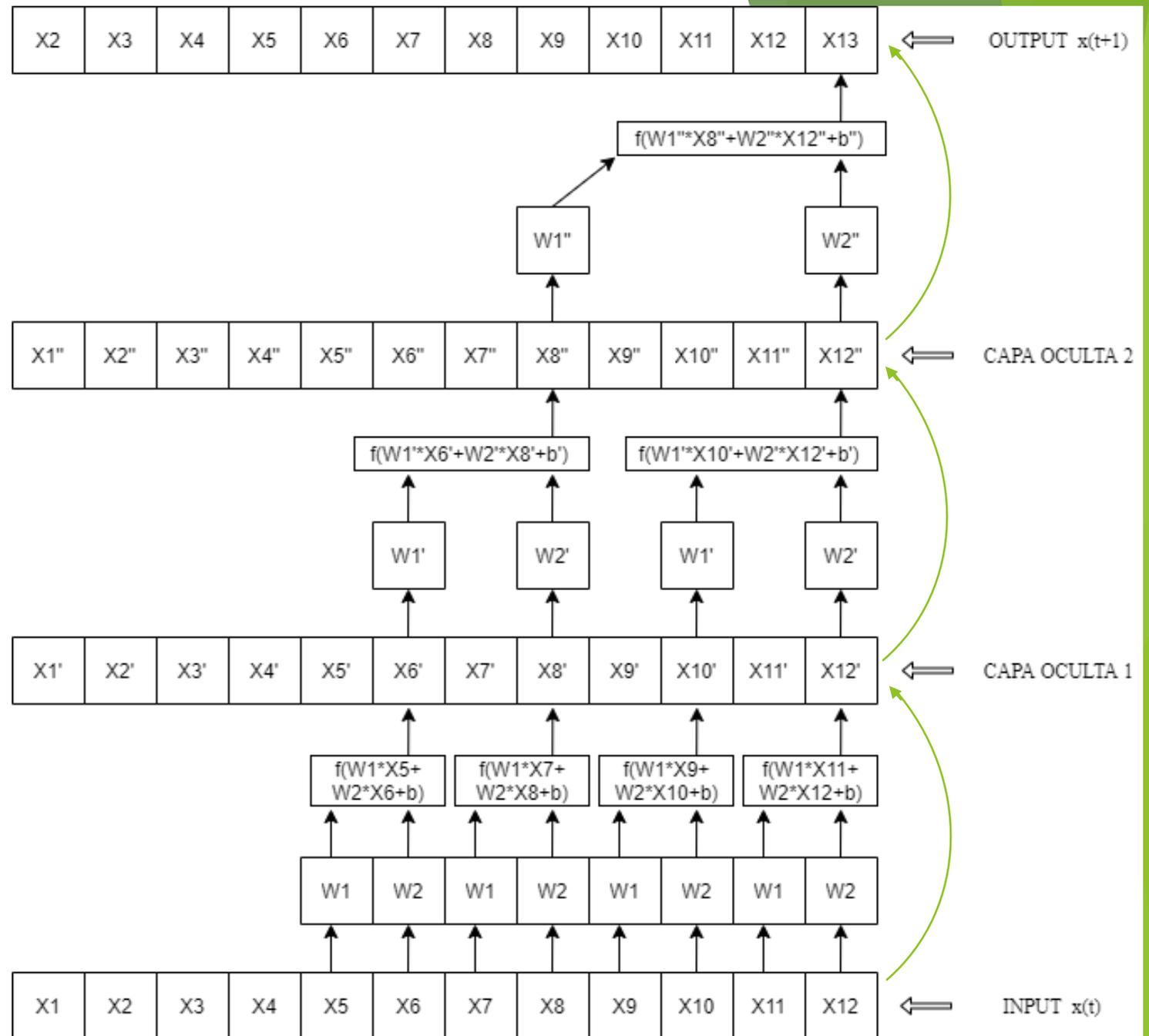
- a) No son causales
- b) Requieren muchas capas



La Arquitectura WaveNet



La Arquitectura WaveNet



Ventajas y Desventajas

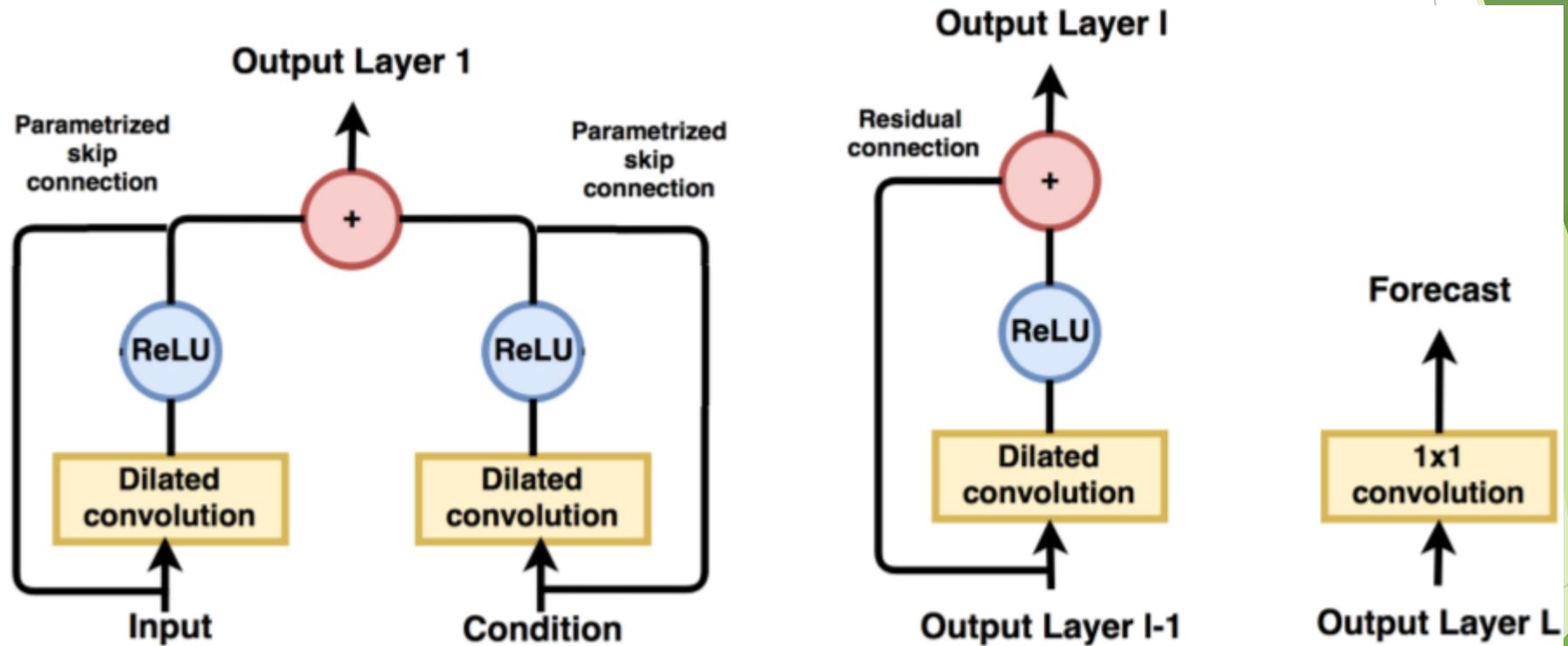
- ✓ Pocos parámetros:

$$R \propto 2^p$$

Donde R es el campo receptivo y p es el número de parámetros. En el modelo no dilatado $R \propto p$.

- ✓ No Lineal
- ✓ Versátil
- ✗ Puede converger a un mínimo local
- ✗ Es un “modelo de caja negra”

WaveNet Condicional



Fuente: Borovykh, et. al. (2018)

EL EXPERIMENTO

The background features abstract, overlapping geometric shapes in various shades of green, ranging from light lime to dark forest green. These shapes are primarily located on the right side of the frame, creating a dynamic, layered effect against the white background.

VARIABLES UTILIZADAS

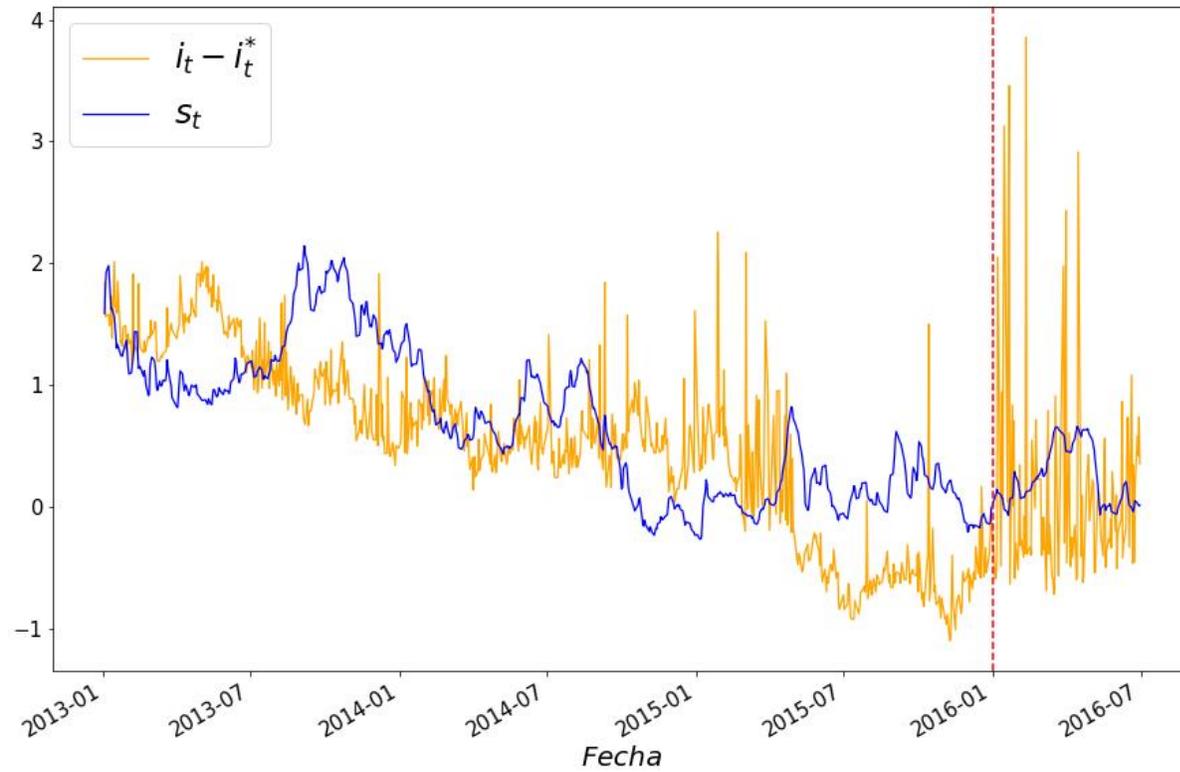
► Frecuencia diaria:

- $s_t = \text{Log}(TC_t)$
- $i_t - i_t^*$
- i_t es la tasa interbancaria, i_t^* es la tasa de Bonos del Tesoro a 10 años.
- El período de calibración es de 2013/01/01 a 2015/12/31
- El período de evaluación es de 2016/01/01 a 2016/06/30

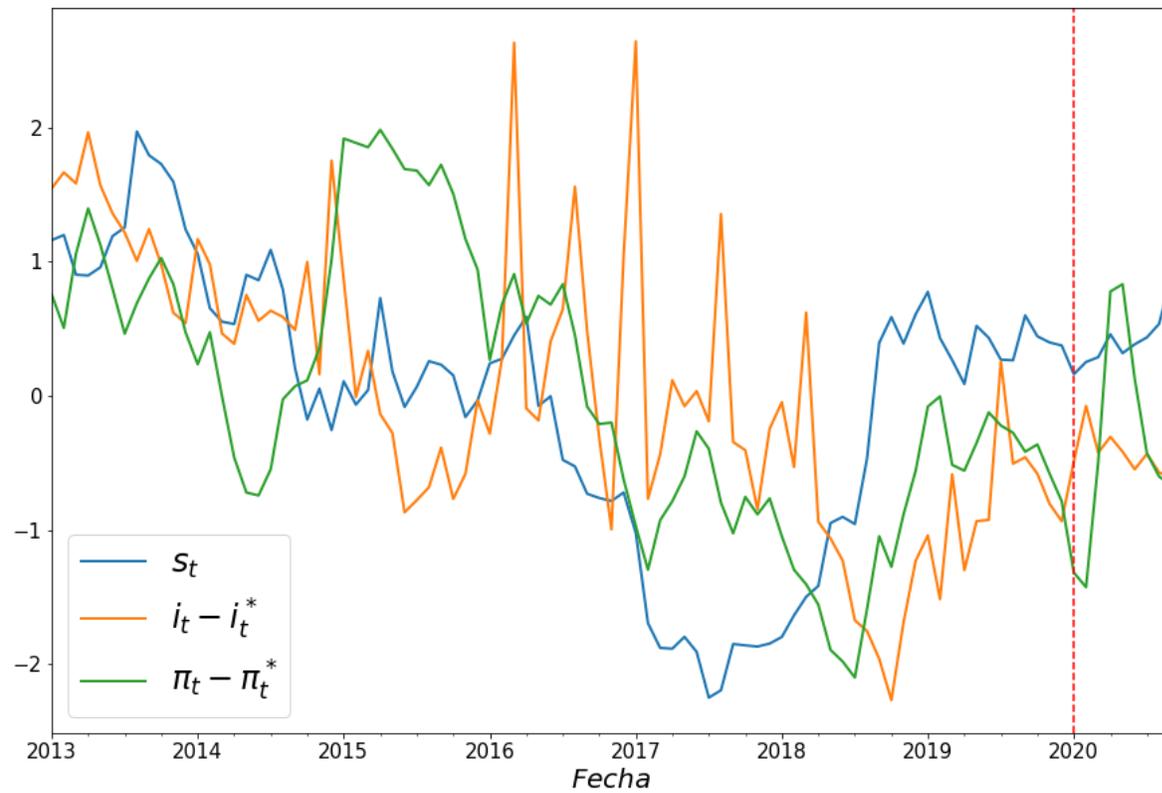
► Frecuencia Mensual:

- $s_t = \text{Log}(TC_t)$
- $i_t - i_t^*$
- $\pi_t - \pi_t^*$
- Para s_t, i_t, i_t^* se utilizaron las últimas observaciones de cada mes. π es la inflación interanual.
- El período de calibración es de 2013/01 a 2019/12
- El período de evaluación es de 2020/01 a 2020/09

Datos Diarios



Datos Mensuales



Modelos *Benchmark*

- ▶ Modelo Caminata Aleatoria (Random Walk):

$$s_{t+1} = s_t + \varepsilon_{t+1}$$

- ▶ Modelo ARIMA(p,d,q):

$$\left(1 - \sum_{i=1}^p \alpha_i L^i\right) (1 - L)^d s_t = \left(1 + \sum_{i=1}^q \beta_i L^i\right) \varepsilon_t$$

- ▶ Modelo ECM de una ecuación:

$$\Delta s_t = \alpha_0 + \alpha_1 s_{t-1} + \beta_1 \Delta s_{t-1} + \sum_i \beta_i x_{t-1}^i + \sum_i \gamma_i \Delta x_t^i + \sum_i \delta_i \Delta x_{t-1}^i$$

Medidas de Bondad de Ajuste (Funciones de pérdida)

$$MSE = \frac{1}{N} \sum_t (\hat{\varepsilon}_t)^2$$

$$MAE = \frac{1}{N} \sum_t |\hat{\varepsilon}_t|$$

$$MAPE = \frac{1}{N} \sum_t \frac{|\hat{\varepsilon}_t|}{|s_t|}$$

$$DA = \frac{1}{N} \sum_t H(\Delta \hat{s}_t \Delta s_t), \quad H(x) = \begin{cases} 1 & x > 0 \\ 0 & x \leq 0 \end{cases}$$

Prueba de Hipótesis

$$PT = \frac{\hat{P} - \hat{P}^*}{\sqrt{\widehat{Var}(\hat{P}) - \widehat{Var}(\hat{P}^*)}} \sim N(0,1)$$

Donde:

$$\hat{P} = DA$$

$$\hat{P}^* = P_x P_y + (1 - P_x)(1 - P_y)$$

$$\widehat{Var}(\hat{P}) = \frac{\hat{P}^*(1 - \hat{P}^*)}{N}$$

$$\widehat{Var}(\hat{P}^*) = (2P_x - 1)^2 q_y + (2P_y - 1)^2 q_x + 4q_x q_y$$

$$P_x = \frac{1}{N} \sum_t H(\Delta \hat{s}_t)$$

$$P_y = \frac{1}{N} \sum_t H(\Delta s_t)$$

$$q_x = \frac{P_x (1 - P_x)}{N}$$

$$q_y = \frac{P_y (1 - P_y)}{N}$$

Especificaciones de Modelos Utilizados

	uWN (d)	cWN (d)	uWN (m)	cWN (m)
Campo receptivo	512	1024	16	48
Parámetros entrenables	46	54	26	42

RESULTADOS Y CONCLUSIONES

Datos Diarios

Adentro de la muestra

MODELO	MSE	MAE	MAPE	DA	PT
RW	0.0031	0.0415	1.0302	0	0
ARIMA	0.0022	0.0357	0.6746	0.6982	10.5961
ECM	0.0022	0.0357	0.6713	0.6769	9.4276
uWN	0.0021	0.0351	0.5995	0.7038	10.9286
cWN	0.0022	0.0356	0.6867	0.6829	9.7641

Afuera de la muestra

MODELO	MSE	MAE	MAPE	DA	PT
RW	0.0019	0.0324	0.7958	0	0
ARIMA	0.0013	0.0267	0.8113	0.7083	4.5928
ECM	0.0013	0.0270	0.7944	0.7583	5.6945
uWN	0.0013	0.0260	0.7474	0.7667	5.8728
cWN	0.1135	0.0758	1.5389	0.7250	4.9741

Datos Mensuales

Adentro de la muestra

MODELO	MSE	MAE	MAPE	DA	PT
RW	0.0817	0.2134	1.0879	0	0
ARIMA	0.0808	0.2117	1.4166	0.5663	1.1201
ECM	0.0757	0.2097	1.3288	0.5854	1.5404
uWN	0.0572	0.1889	0.7378	0.7229	3.9953
cWN	0.0380	0.1465	0.6679	0.7590	4.8784

Afuera de la muestra

MODELO	MSE	MAE	MAPE	DA	PT
RW	0.0269	0.1353	0.3880	0	0
ARIMA	0.0270	0.1371	0.3979	0.6667	0.6013
ECM	0.0253	0.1349	0.3785	0.6667	1.52134
uWN	0.0530	0.1742	0.3869	0.3333	0.6013
cWN	0.0916	0.2617	0.6471	0.2222	-1.1364

Conclusiones

- ▶ Los modelos WaveNet presentan el mejor ajuste en la mayoría de los casos, a excepción del período afuera de la muestra para los datos mensuales.
- ▶ En los mismos tres casos, los modelos WaveNet contienen información sobre la dirección del cambio, lo cual los diferencia del modelo de caminata aleatoria.

Para futuros proyectos

- ▶ Realizar pronóstico con un horizonte $h > 1$, para poder realizar prueba de Diebold-Mariano.
- ▶ Entrenar modelo con una función de pérdida diferente

MUCHAS GRACIAS