

RNN, LSTM, GRU para forecasting

Forecast de demanda, ventas y consumo eléctrico.

CONCEPTOS CLAVE

- RNN: $h_t = \tanh(W_x x_t + W_h h_{t-1} + b)$. Mismos pesos cada paso.
- Problema: gradientes desaparecidos o explotados con secuencias largas.
- **LSTM** resuelve con tres puertas + memoria C_t ; **GRU** simplifica a dos puertas.
- En 2026: empieza por GRU (menos parámetros). Si no basta, sube a LSTM o Transformer (Día 09).

EJEMPLO A MANO

Tres pasos a mano de una RNN sobre demanda escalada ($W_x = 0,01$, $W_h = 0,5$, $b = -0,2$, $w_y = 1,5$):

Paso	x_t	h_{t-1}	$h_t = \tanh(\cdot)$	y_t
1	100	0,00	$\tanh(0,80) = 0,66$	0,99
2	108	0,66	$\tanh(1,21) = 0,84$	1,26
3	112	0,84	$\tanh(1,34) = 0,87$	1,30

El estado oculto h acumula información del pasado y "predice" la siguiente venta.

CONCEPTO → APLICACIÓN A FORECASTING

Caso	Variable predicha	KPI
Demanda retail (Inditex)	Unidades/SKU/semana.	MAPE, días sin stock.
Consumo eléctrico (Iberdrola)	MWh/hora/región.	RMSE €/MWh.
Tráfico datos (Telefónica)	GB/celda/5 min.	Pico no servido.
Precio energía (Endesa)	€/MWh hora siguiente.	€ de cobertura.
ASR (atención cliente)	Texto token siguiente.	WER, tiempo manual.

Idea clave del Día 06. LSTM y GRU ganan a ARIMA **solo cuando hay covariables externas** (promoción, precio competencia, calendario). Si la serie es estacionaria pura, ARIMA es difícil de batir.

Para llevarte: lectura Olah, *Understanding LSTM Networks*; práctica: LSTM 2 capas + dropout 0,2 sobre Online Retail II.