

RNN para Sunspots (SIDC)

Descripción del dataset

El dataset usado es el **Daily total sunspot number (v2.0)** del **WDC-SILSO, Royal Observatory of Belgium** ([enlace](#)), con datos desde **1818-01-01** hasta la fecha más reciente.

Incluye:

- Año, mes, día y fecha fraccionaria.
- Número diario de manchas solares (-1 = dato faltante).
- Desviación estándar diaria.
- Número de observaciones.
- Indicador de valor definitivo (1) o provisional (0).

Capa investigada – torch.nn.RNN

Input esperado. Tensor con forma (batch, seq_len, input_size) usando batch_first=True.

Output.

- output: secuencia de estados ocultos con forma (batch, seq_len, D*hidden_size), donde D=1 al ser unidireccional.
- h_n: último estado por capa y dirección con forma (num_layers*D, batch, hidden_size).

Parámetros usados. input_size=1, hidden_size=16, num_layers=1, batch_first=True.

Metodología

- **Ventanas** se usaron las longitudes 5, 10, 20 y 100.
- **Modelo:** nn.RNN seguida de una capa lineal de salida.
- **Entrenamiento:** pérdida MSE, optimizador Adam (1e-2), 50 épocas.
- **Preparación de datos:** partición temporal en train/valid/test; visualización de la serie diaria y su promedio mensual antes del modelado.

Resultados

		Train MSE	Val MSE	Test MSE	Tiempo	grad_norm
Ventana	seq_len	(ep50)	(ep50)			(ep50)
small	5	0.002734	0.001850	0.001200	0.70 s	0.055
initial	10	0.002636	0.001834	0.001156	1.37 s	0.016

		Train MSE	Val MSE			grad_norm
Ventana seq_len		(ep50)	(ep50)	Test MSE	Tiempo	(ep50)
medium	20	0.002693	0.001844	0.001076	2.89 s	0.040
large	100	0.008931	0.008487	0.006062	15.23 s	0.057

- **Ventana 20 → mejor desempeño:** menor error en test (0.001076), entrenamiento estable y buen balance entre contexto, precisión y costo.
- **Ventana 100 → peor desempeño:** error alto (0.006062) y mayor tiempo (15.23 s), afectada por desvanecimiento del gradiente y limitaciones de la RNN simple.
- **Ventanas 5 y 10 → aceptables,** con menor costo de tiempo, aunque menos precisas que 20.

El rango óptimo está entre **10–20 pasos**.

Conclusiones

1. Se confirmó que usar ventanas intermedias (10–20 pasos) ofrece el mejor compromiso entre capturar la dinámica de las manchas solares y mantener un error bajo. Ventanas demasiado largas (100 pasos) degradan la predicción y aumentan el costo de entrenamiento.
2. El modelo evidencia el problema de vanishing gradient en secuencias extensas, lo que provoca pérdida de amplitud en las predicciones. Esto sugiere que arquitecturas más avanzadas (LSTM, GRU) serían más adecuadas para dependencias de largo plazo.
3. El enfoque many-to-one, en el que cada ventana genera una sola predicción, resultó suficiente para el fenómeno estudiado, sin requerir bidireccionalidad ni regularización adicional.
4. Los resultados empíricos coinciden con lo discutido en clase: la escala característica del ciclo (~11 años) se captura mejor con longitudes de ventana intermedias, mientras que secuencias muy largas introducen atenuación de la señal.
5. No se observaron exploding gradients.

Referencias

- [RNN Documentation \(PyTorch\)](#)
- [LSTM Documentation \(PyTorch\)](#)
- [Dataset Sunspots \(WDC-SILSO\)](#)
- [Información detallada del dataset diario \(SILSO\)](#)