

Feed Forward NN (Red Simple)

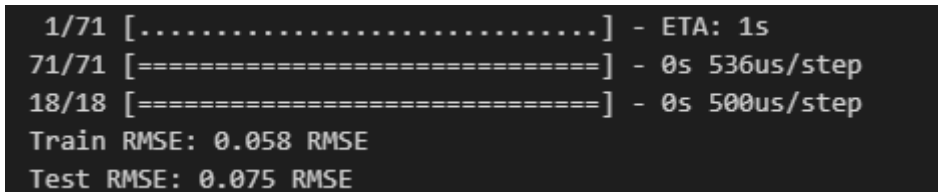
Las redes neuronales feedforward, también conocidas como FFNN (Feedforward Neural Networks), son uno de los tipos más básicos y tradicionales de redes neuronales. A diferencia de las redes neuronales recurrentes, que tienen conexiones cíclicas, las FFNN fluyen en una sola dirección: desde la entrada hasta la salida, sin ciclos.

Estas redes son especialmente útiles para tareas donde los datos pueden considerarse independientes y no se necesita capturar secuencias o patrones temporales, a diferencia de lo que ocurre con las RNN. Por su simplicidad y estructura directa, son ideales para problemas de clasificación y regresión en muchos campos, desde finanzas hasta medicina.

Para el modelo FFNN en cuestión, se utiliza una capa oculta compuesta por 10 unidades o neuronas. La entrada a la red es un vector de características con una longitud definida, en este contexto, basada en el número de `time_steps` que se consideren relevantes para predecir el valor siguiente de manchas solares. La función de activación 'ReLU' es utilizada en la capa oculta para introducir no linealidad en la red, permitiendo así que la red pueda aprender relaciones más complejas en los datos. Por otro lado, la capa de salida utiliza una función de activación lineal, adecuada para problemas de regresión como la predicción de manchas solares.

Resultados

Luego de entrenar el modelo y realizar predicciones se encontraron los valores de RMSE como se muestra en la siguiente figura.



```
1/71 [.....] - ETA: 1s
71/71 [=====] - 0s 536us/step
18/18 [=====] - 0s 500us/step
Train RMSE: 0.058 RMSE
Test RMSE: 0.075 RMSE
```

Fig. 1. RMSE de red FFNN

Los resultados obtenidos tras entrenar la red FFNN son notablemente prometedores. A partir de los valores del error cuadrático medio (RMSE) que obtuvimos:

Train RMSE: 0.058 y Test RMSE: 0.075 es evidente que la red ha logrado capturar las tendencias y patrones subyacentes del conjunto de datos con una precisión considerable.

Estos errores tan bajos indican que la red FFNN ha sido capaz de realizar predicciones muy cercanas a los valores reales, tanto en el conjunto de entrenamiento como en el conjunto de prueba. La diferencia entre los errores de entrenamiento y prueba es mínima, lo que sugiere que el modelo ha generalizado bien y no ha preajustado al conjunto de entrenamiento.

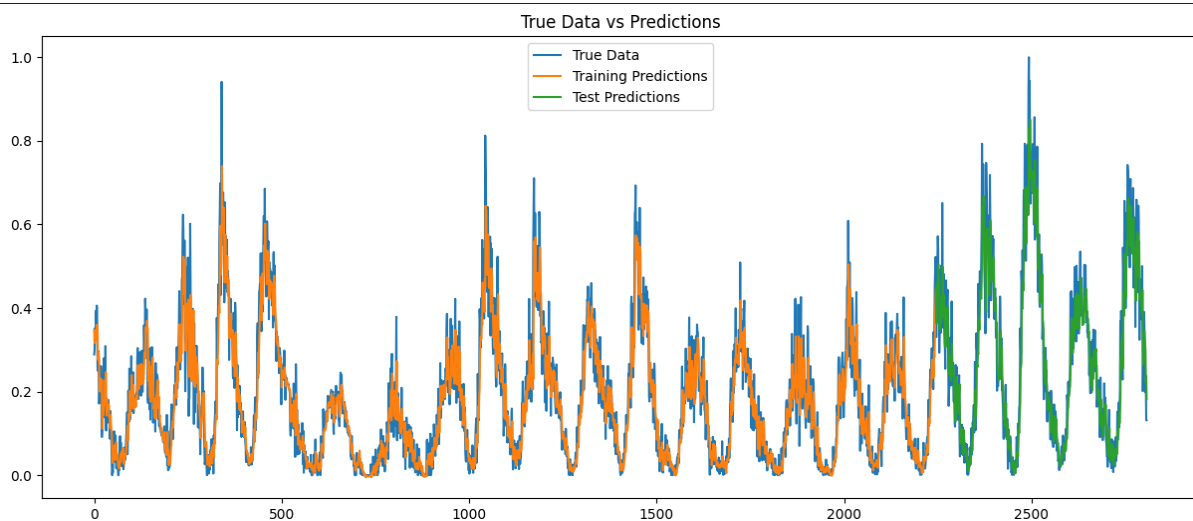


Fig. 2. Gráfica de predicción vs. valores reales FFNN

Pros y Contras

Pros	Contras
Las FFNN tienen una estructura directa y sencilla, lo que las hace fácilmente comprensibles e implementables.	las FFNN no tienen memoria de estados anteriores, lo que las hace inadecuadas para datos secuenciales o temporales.
Son aptas para una amplia variedad de tareas, desde clasificación hasta regresión, en diversos campos.	Pueden ser propensas al sobreajuste, especialmente cuando tienen muchas capas o neuronas y pocos datos de entrenamiento.
No tienen conexiones cíclicas o recurrencias, lo que puede hacer que su entrenamiento y predicción sean más rápidos en comparación con arquitecturas más complejas.	Sin estructuras especializadas (como convoluciones o recurrencias), pueden requerir más datos o ingeniería de características para capturar patrones complejos.

RNN

Las redes neuronales recurrentes son un tipo de arquitectura de redes neuronales especialmente diseñadas para trabajar con datos secuenciales, lo cual lo hace ideal para analizar series de tiempo como el dataset utilizado en este laboratorio. Este tipo de red tiene conexiones internas que permiten mantener una cierta memoria de entradas anteriores al momento de procesar nuevas entradas, lo que le hace poder capturar patrones y dependencia a lo largo del tiempo en los datos. En concreto para esta ocasión se utilizó esta red para predecir valores futuros de las manchas solares basándose en observaciones pasadas en el tiempo (por mes).

Para esta red en particular se utilizan 10 unidades/neuronas en la capa oculta y 1 unidad densa la cual representa la capa de salida. El tipo de entrada es una secuencia de dimensión de tiempo `time_steps` y una dimensión (representando el valor de manchas solares), adicionalmente, la red cuenta con una función de activación de tangente hiperbólica para la capa oculta al igual que en la capa de salida.

Resultados

Luego de entrenar el modelo y realizar predicciones se encontraron los valores de RMSE como se muestra en la siguiente figura.

```
6/6 [=====] - 0s 1ms/step
2/2 [=====] - 0s 2ms/step
Train RMSE: 0.056 RMSE
Test RMSE: 0.081 RMSE
```

Fig. 3. RMSE de red RNN

Los valores del error cuadrático medio encontrados nos demuestran que la red esta haciendo un muy buen trabajo prediciendo los valores del conjunto de entrenamiento y de prueba, los valores se encuentran dentro de un rango de diferencia que para este conjunto de datos puede ser insignificante, por lo que se le podría considerar a esta red RNN como un buen acercamiento a este tipo de conjunto de datos debido a su precisión tan alta. De una manera más visual, en la siguiente figura se puede observar las predicciones (línea naranja) vs los datos reales (línea azul).

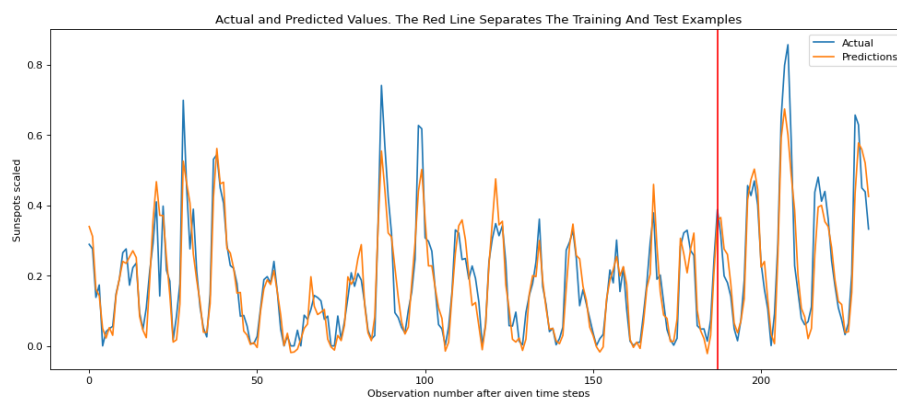


Fig. 4. Gráfica de predicción vs. valores reales RNN

Pros y Contras

Pros	Contras
Las RNN son adecuadas para modelar datos secuenciales como series de tiempo y lenguaje natural.	Puede sufrir del problema de desvanecimiento y explosion del gradiente.

Pueden manejar secuencias de longitud variable, lo cual es útil cuando las secuencias tienen diferentes longitudes.	Son más lentas a comparación de otras arquitecturas debido a su recurrencia, ya que el cálculo en cada paso de tiempo depende del tiempo anterior
Son eficientes en tareas de predicción y generación de secuencias futuras a partir de los datos de entrada iniciales.	Debido a su naturaleza secuencial son menos eficientes en términos de paralelismo durante el entrenamiento.

LSTM

Las redes Long Short-Term Memory son una variante de las redes neuronales recurrentes (RNN) diseñadas para abordar el problema del desvanecimiento del gradiente y mejorar el modelado de dependencias a largo plazo en series de tiempo u otras secuencias temporales. Su diseño incorpora unidades de memoria especializadas que permiten retener y gestionar información relevante a lo largo de intervalos temporales más extensos. Estas características mencionadas lo hacen ideal para realizar una red neuronal que ayude en las predicciones de las manchas solares a través del tiempo.

La red diseñada en particular para este laboratorio cuenta con 50 unidades en la capa LSTM con una función de activación 'relu'. Esta capa procesa la secuencia y captura patrones temporales complejos encontrados en el conjunto de datos dado. Además, cuenta con una capa densa totalmente conectada la cual produce la salida final de la red, no contiene ninguna función de activación.

Resultados

A continuación, luego de su respectivo entrenamiento se muestra en la figura los valores del RMSE para esta red.

```
71/71 [=====] - 0s 1ms/step
18/18 [=====] - 0s 1ms/step
Train RMSE: 0.058 RMSE
Test RMSE: 0.076 RMSE
```

Fig. 5. RMSE de red LSTM

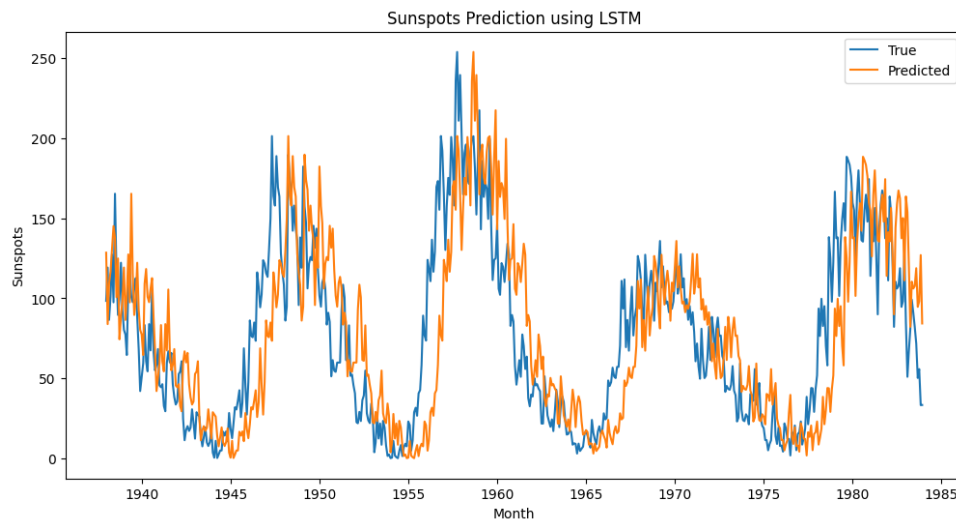


Fig. 6. Gráfica de predicción vs. valores reales LSTM

Pros y Contras

Pros	Contras
Pueden mantener la información relevante al modelo y conjunto de datos a través de intervalos temporales extensos.	Son computacionalmente intensivas a comparación de otras redes neuronales mas simples debido a la mayor cantidad de parámetros y operaciones que realizan en cada intervalo de tiempo.
Prevención del desvanecimiento del gradiente a través de las unidades de memoria y las compuertas.	Es sensible a que el modelo memorice valores atípicos, lo que puede llevar a un sobreajuste en ciertos casos.
Eficaz al momento de modelar patrones complejos en series de tiempo, incluyendo tendencias, estacionalidades y relaciones no lineales.	Requiere de extensiva configuración y ajuste de los hiperparametros para asi encontrar los ajustes óptimos para la red y el conjunto de datos.

Red Elegida y Justificación

La red elegida basada en los resultados observados es la red que utiliza **LSTM** esto se debe principalmente a la exactitud que ofrecen sus predicciones para este conjunto de datos en concreto, aunque tenga una exactitud similar a la de la red RNN, el LSTM al ser una modificación posterior, tiene muchas ventajas que la hacen superior a la RNN y la hace la más adecuada para para resolver el problema. Principalmente esta respuesta se debe a que si se quisiera utilizar con un conjunto mucho mas extenso para asi expandir el rango del aprendizaje de la red, la red LSTM es la ideal, ya que a diferencia de las RNN, las redes LSTM pueden capturar dependencias a largo plazo en secuencias, lo cual es crucial en

situaciones en las que la relación entre elementos en la secuencia esta separada por varios pasos de tiempo. Adicionalmente, las redes RNN tienden a sufrir de problemas de desvanecimiento y explosion del gradiente, lo que dificulta el entrenamiento de modelos en secuencias largas, las LSTM utilizan compuertas que controlan el flujo de la información y ayudan en la reducción de este tipo de problemas. Aunque la red Feed Forward NN y la red RNN puedan ofrecer cierta eficiencia computacional, lo que haga más fácil su utilización en máquinas con recursos limitados estos tienden a tener fallos mientras más crezca el conjunto de datos y este al mismo tiempo se vuelva más complejo, lo que realmente no vale la pena si se planea utilizar la red a largo plazo. Por lo que por su longevidad, exactitud y eficiencia la red LSTM se puede considerar como la más óptima para la resolución de este problema.