

Hiperparámetros

- Total de redes neuronales desplegadas: 2

Tabla

Nombre	Espacio de búsqueda	Valor definido	Razón
Batch size	[16,32,64]	32	Mejor rendimiento, recomendado en varios sitios inc. tensorflow/keras docs
Shuffle dataset	[True,False]	False	Ideal para LSTM porque tienen memoria
Data standarization	[True,False]	True	Debido al uso de Tanh & mejor MSE
Timestep size	[1s,1min,1hr,1day]	1hr	Tamaño del dataset, usando sólo días son muy pocos datos
Window 'past' size (LSTM input size)	[1 month,1 week,1 day]	1 month	
Window 'future' size (LSTM output size)	[1 hour, 1 day, 1 week]	1 week	
Window shift	[1 month,1 week, 1 day]	1 week	
max_epochs	[10,20,25,30,50,100,200]	30	Va más a underfitted, mayores valores sobreajustan demasiado rápido
Learning rate	[0.0001,0.001,0.01,0.1]	0.0001	underfitted pero da mejores resultados que el resto
Optimizer	[Adam]	Adam	Usado siempre reemplazando al SGD explícitamente, Muestran SGD como obsoleto
Kernel initializer	[tf.zeros,default]	default	En los docs de tensorflow lo recomiendan como zeros, pero eso hace que nunca entrene en este caso
Early Stopping	[True]	True	
Early Stopping Patience	[2,3,5]	3	
Loss function	[MSE]	MSE	
Evaluation Metrics	[MSE,RMSE]	[MSE,RMSE]	
Activation function	[ReLU, Tanh]	Tanh	Obligatorio para Usar CUDA
LSTM units	[16, 32,64,128]	32	Mejor rendimiento entre todas

Arquitectura

RNN 1

- 1 capa LSTM | 32 unidades | tanh -> mejores resultados
- 1 capa densa | $24 \times 7 \rightarrow 168$ | linear -> una por hora de la semana de salida
- 1 capa de reshape | $2471 \rightarrow 168 \times 1$ | básicamente una capa de flatten, el *1 es el número de features, en este caso sólo hay 1, originalmente se planeaba retornar [open,high,low,close]

RNN 2

- 1 capa GRU | 32 unidades | tanh -> mejores resultados <- Esta entrena más rápido que LSTM y acaba antes por el early stopping, keras parece ser que es más eficiente en cuanto a recursos y no tiene mecanismo de 'memoria'
- 1 capa densa | igual al LSTM
- 1 capa de salida | igual al LSTM