Trabajo Práctico 7

Camussoni Francesco[1,*]

[1]Instituto Balseiro, Universidad Nacional de cuyo
[*]camussonif@gmail.com

ABSTRACT

Se implementó una red neuronal recurrente con caracteristicas LSTM para predecir la cantidad de pasajeros de avión según la fecha del dataset airline-passengers. Finalmente, se comparo su rendimiento con una red neuronal recurrente de capas densas.

LSTM

Se utilizo el set de datos de airline-passengers el cual establece la cantidad de pasajeros que viajaron en avión según determinado mes de cierto año como se muestra en la **Figura 4**.

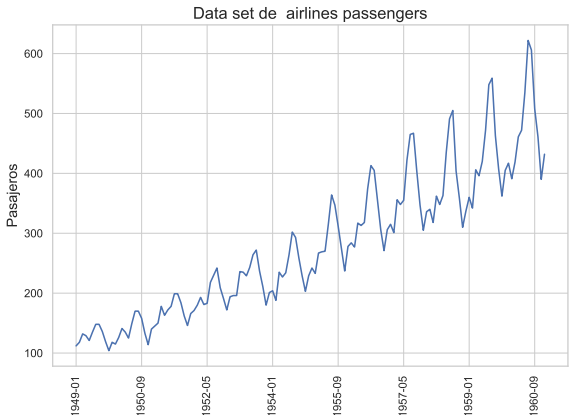

Figure 1. Datos originales.

En primera instancia se normalizaron los datos entre 0 y 1 y se reformatearon los datos de la siguiente manera:

$$\begin{cases} X = (x(t), x(t+1), ..., x(t+l-1)) \\ Y = x(t+l) \end{cases}$$

En donde l es un entero positivo que determina la retrospección de los datos. Luego, se agrego un ruido con media 0 y dispersión de 0.2 al vector Y.

A continuación, se separo el set de datos en training y test con un 50% de los datos originales en cada caso. Finalmente, en cuanto a preprocesado de datos, se modificó la dimension de los datos para que puedan ser utilizados por la red LSTM que requiere el siguiente tipo de input: [batch, timesteps, feature]. 'batch' no es necesario identificarlo a priori y timesteps=1 en nuestro caso. 'feature' será iguala la retrospección elegida.

Luego, se definió la siguiente arquitectura:

```
Model: "red_recurrente_LSTM"
_________________________________________________________________Layer (type)                 Output Sha
=================================================================LSTM (LSTM)                  (None, 16
_________________________________________________________________Dense (Dense)                (None, 1)
=================================================================Total params: 1,169
```

```
Trainable params: 1,169
Non-trainable params: 0
```

En donde los parámetros se dejaron por defecto activation="tanh", recurrent_activation="sigmoid", kernel_initializer="glorot_uniform", recurrent_initializer="orthogonal", bias_initializer="zeros",

Así, se entreno a la red neuronal por 200 épocas, un learning rate de 1e-3, una retrospección de 1 y MSE como función de costo en donde se obtuvieron los resultados que se muestran en las **Figuras** 2 y 3.

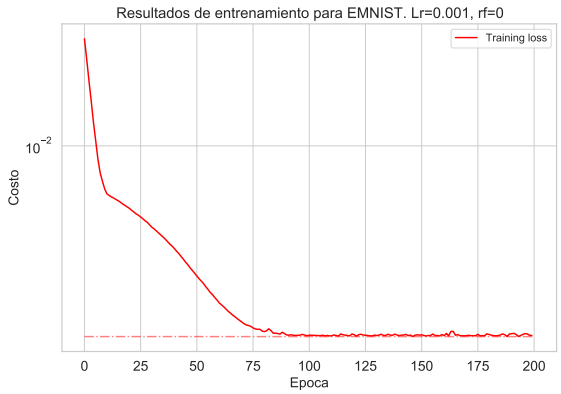

Figure 2. Costo para los datos de training según la época.

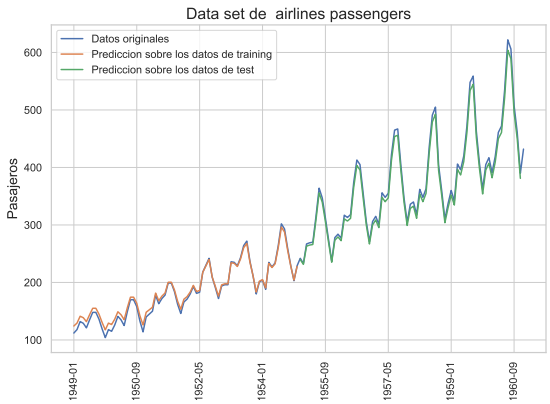

Figure 3. Datos originales y predicción sobre los datos de training y sobre los datos de test.

En donde se observa un gran rendimiento en la predicción con un error de 0.0015 en los datos de training y 0.0058 en los datos de test según MSE

A continuación se realizó el mismo entrenamiento modificando la retrospección de 1 a 25 en donde en cada instancia se obtuvo el error en los datos de test obteniendo el resultado que se muestra en la **Figura ??**.

Figure 4. Loss sobre los datos de test en función de la retrospección.

En donde no se observa un comportamiento definido, si no más bien un comportamiento azaroso o de ruido. Probablemente el la elección de la retrospección no tenga influencia en el loss de test pero si en las predicciones a futuro o Forecasting

Capas densas

Finalmente, se implementó una arquitectura similar en donde se remplazó la capa LSTM por una capa densa de la misma cantidad de neuronas con activación relu en donde se obtuvieron los resultados que se muestran en las **Figuras** 5 y 6.

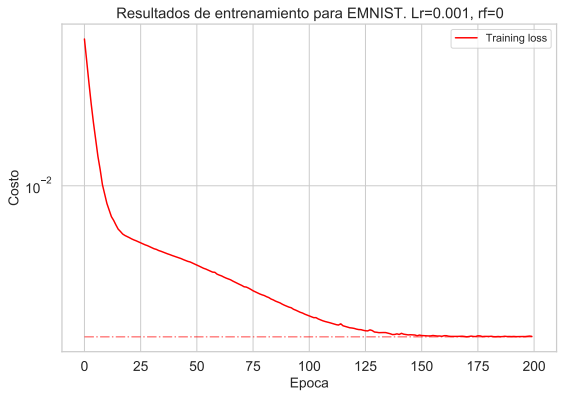

Figure 5. Costo para los datos de training según la época.

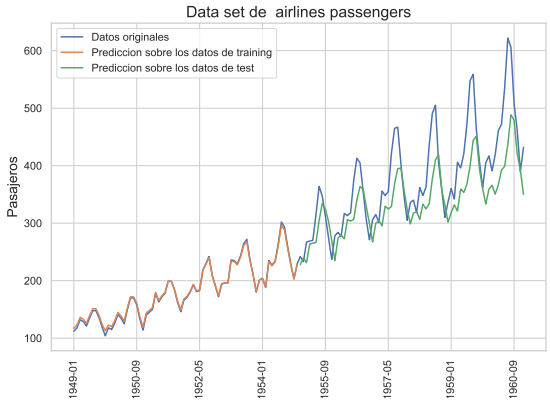

Figure 6. Datos originales y predicción sobre los datos de training y sobre los datos de test.

En donde se observa un rendimiento menor que al utilizar capas LSTM. Esto es debido, como se explicó anteriormente, a que las mismas están bien diseñadas para la prediccion de información dependiente del tiempo, dado que puede haber cierto lag de duración desconocida entre los eventos importantes o relevantes de la información en donde la capa LSTM realizá un balance olvidar parte de lo aprendido e incorporar la nueva información, mientras que las capas densas no tienen esta propiedad.