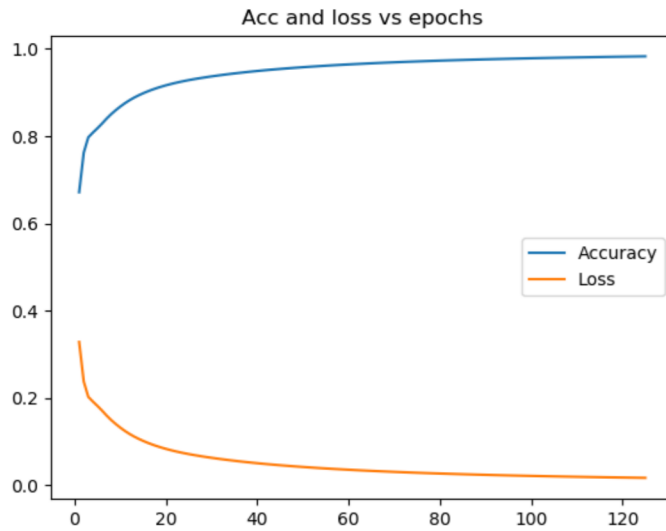


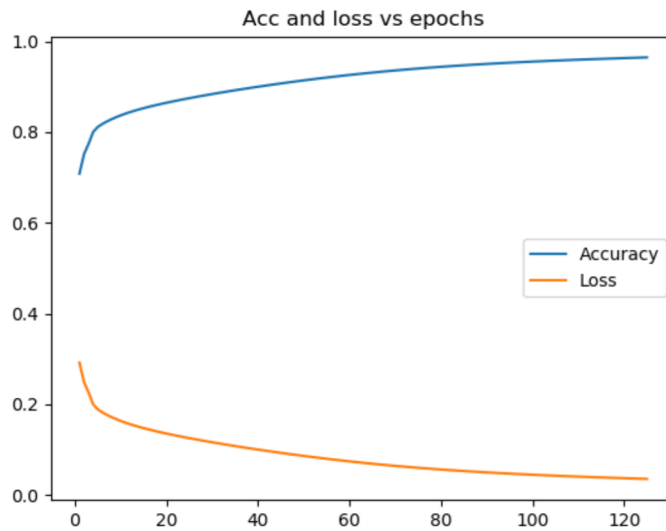
- Pruebas utilizando tres diferentes proporciones de train y test:
En estas pruebas utilizamos un lr de 0.01 y 125 epochs:
 - Usando 30% de test:



Accuracy del modelo con la parte de test: 96.35%

Pérdida del modelo con la parte de test: 0.04

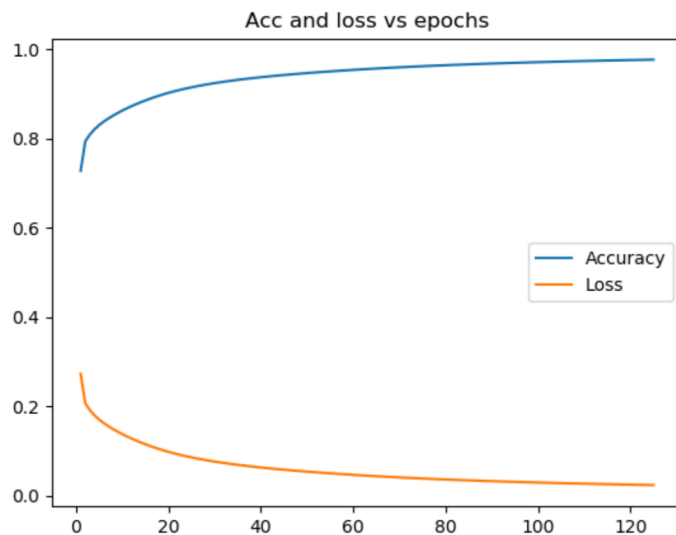
- Usando 20% de test:



Accuracy del modelo con la parte de test: 92%

Pérdida del modelo con la parte de test: 0.08

- Usando 40% de test:



Accuracy del modelo con la parte de test: 95.3%

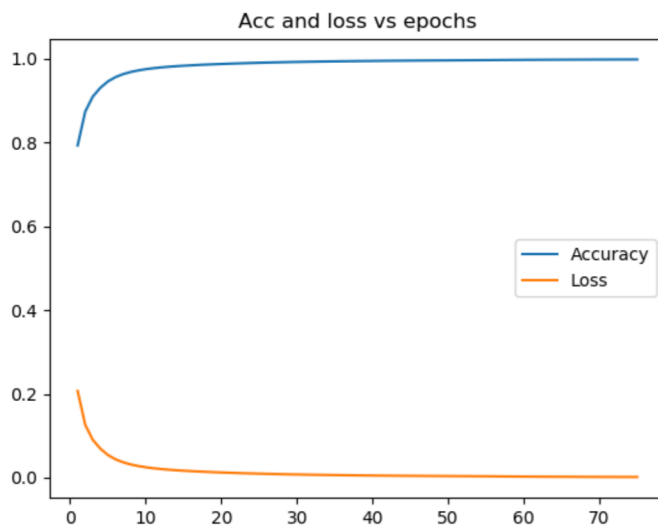
Pérdida del modelo con la parte de test: 0.047

Como podemos ver, en estas pruebas la mejor división de la dataset es el 30% de test aproximadamente. Ya que usado 20% de test ya empieza a entrar a overfitting y con 40% ya entra en underfitting.

- Cambio de los hiperparametros:

En nuestro caso vamos a cambiar el número de epochs y el learning rate.

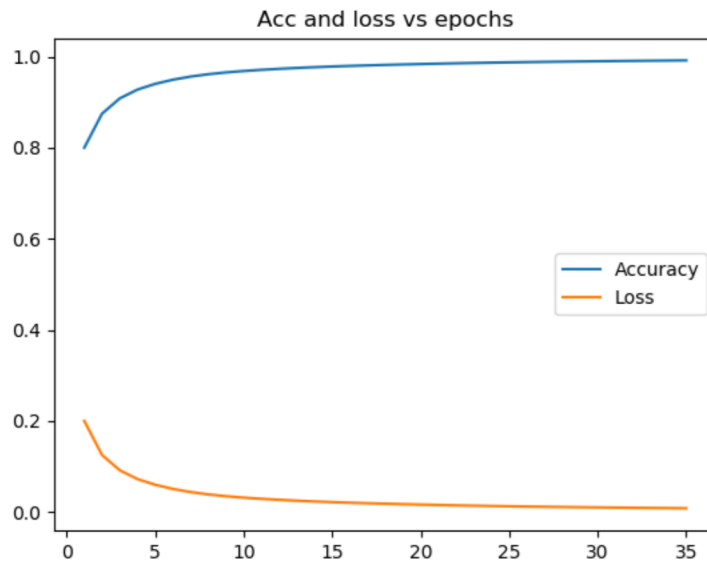
- lr de 0.1, 75 epochs y un test del 30%:



Accuracy del modelo con la parte de test: 95.78%

Pérdida del modelo con la parte de test: 0.042

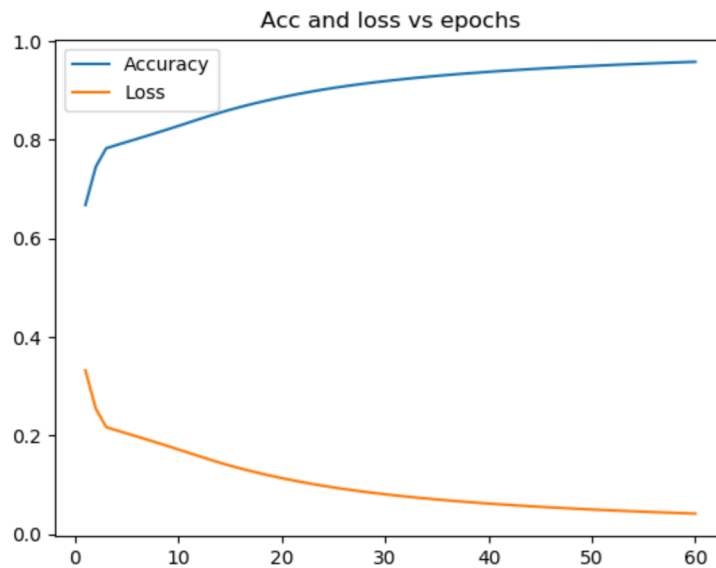
- lr de 0.1, 35 epochs y un test del 30%:



Accuracy del modelo con la parte de test: 96.63%

Pérdida del modelo con la parte de test: 0.033

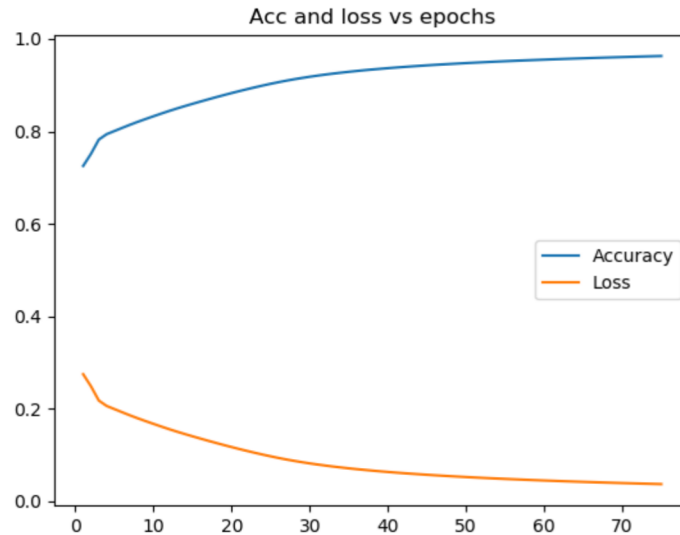
- lr de 0.01, 60 epochs y un test del 30%:



Accuracy del modelo con la parte de test: 93.9%

Pérdida del modelo con la parte de test: 0.06

- lr de 0.01, 60 epochs y un test del 30%:



Accuracy del modelo con la parte de test: 94.6%

Pérdida del modelo con la parte de test: 0.053

- Análisis:

Como podemos ver nuestro modelo no necesita un learning rate tan pequeño como tampoco necesita tantas epochs, ya el modelo con el lr más grande y menos epochs es el que tuvo mayor acc en la parte del test, tomando en cuenta que siempre la dataset es dividida de manera aleatoria, además de que en las gráficas de comportamiento vemos que no hay picos ni ningún comportamiento extraño, dándonos a entender que no debemos aprender tan lento para poder encontrar el mínimo global, declarando que los epochs que vamos a usar son 35 con un lr de 0.1, dándonos el mejor comportamiento. Usando un leer más pequeño también logramos aprender de manera satisfactoria, pero necesitamos muchas más epochs y caer más fácil en overfitting, como también en underfitting.