Homework 1

Yusheng Zhao

<2023-10-01 Sun>

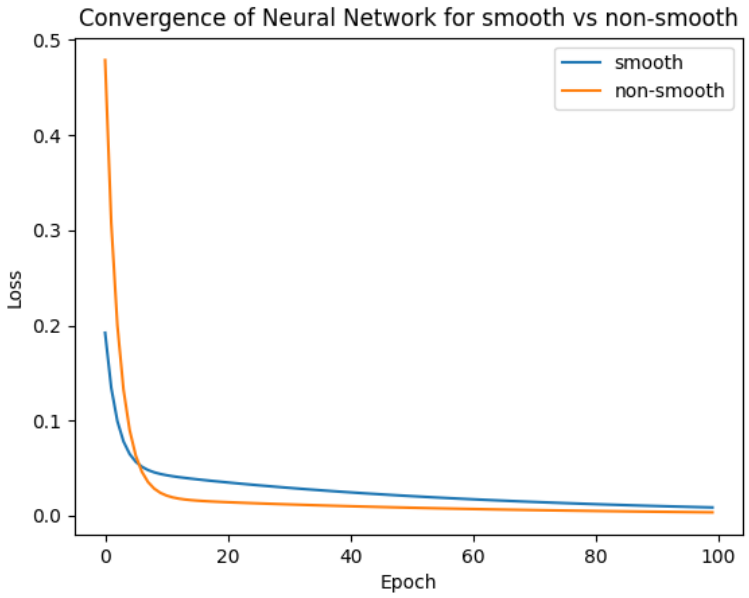
1 Answer 2

In this problem, I will investigate the influence of number of epochs a two layer neural is trained on the accuracy of the model. I investigate for five different types of functions.

1.1 Smooth vs Non-smooth function

For the smooth function, I use the function $f(x) = x$, and for the non-smooth function, I use the function $f(x) = x - \lfloor x \rfloor$.

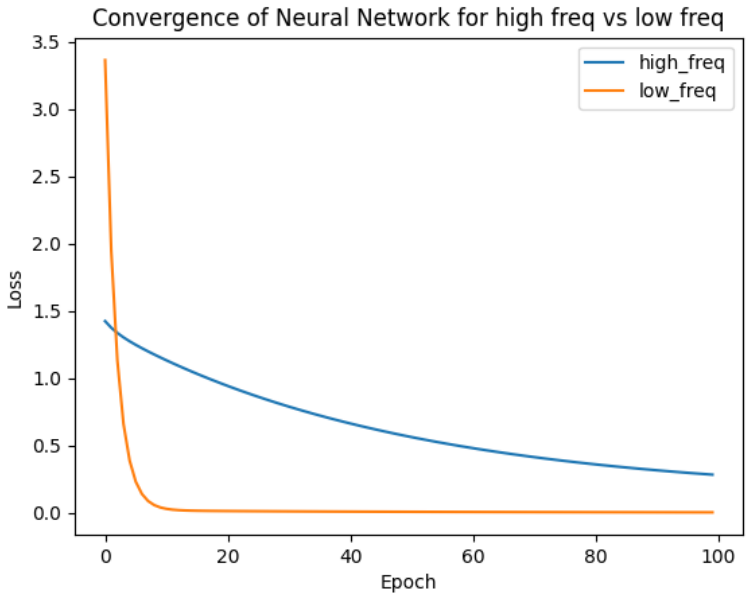
As indicated in the graph below, the non-smooth is more difficult to train than the smooth function. I.e, it reaches low loss in later epochs. The accuracy of the model for the non-smooth function also starts out to be worse than that of the smooth function.



1.2 High Freq vs Low Freq

For the high frequency function, I use the function $f(x) = \sin(5x) + cos(5x)$, and for the low frequency function, I use the function $f(x) = \sin(x) + cos(x)$.

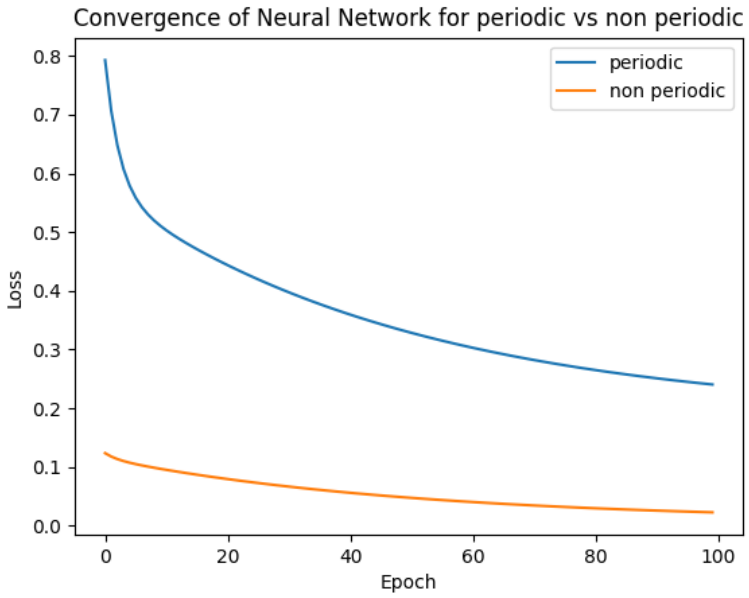
As indicated in the graph below, the high frequency is more difficult to train as it taks more epochs to stabilize.



1.3 Periodic vs Non periodic function

For the periodic function, I use the function $f(x) = sin(2 * \pi * x)$, and for the non periodic function, I used $f(x) = x^2$

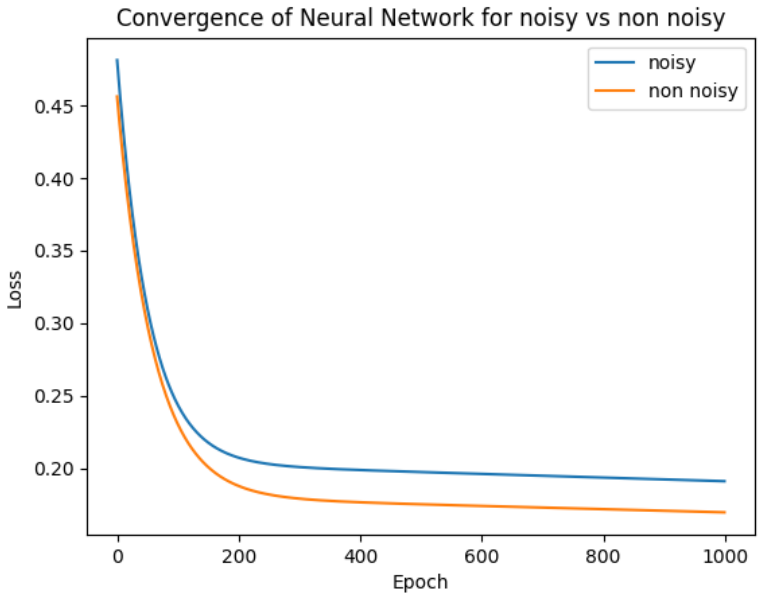
As indicated in the graph below, the non periodic function is easier to train.



1.4 Noisy vs Non Noisy function

For the noisy funciton, I used $f(x) = sin(2 * \pi * x)$ plus some random noise. I used $f(x) = sin(2 * \pi * x)$ for noiseless one.

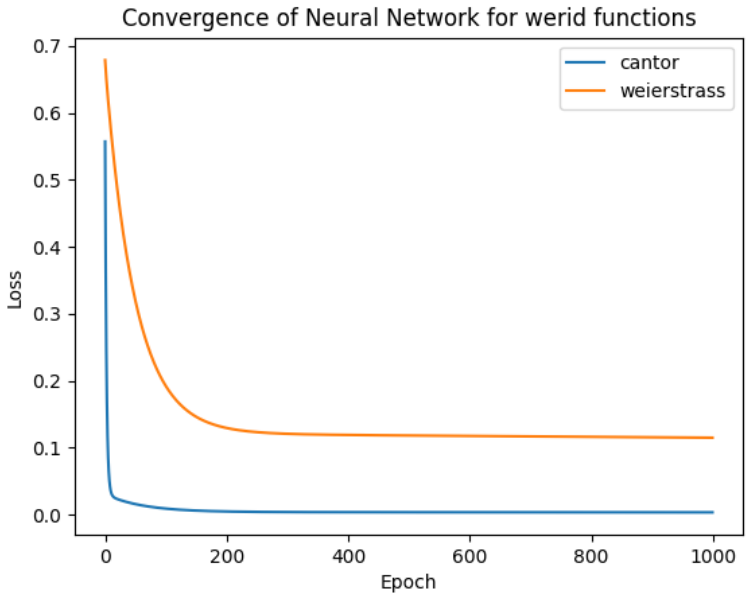
It is unclear who converges in less epochs. During the multiple experiment I carried out, I observe sometimes one function converges faster while sometimes the other one converges faster. However, they both take considerably longer to converge.



1.5 Cantor vs Weirestress function

For the last group, I compare training epochs required to approach cantor and weirestress function.

It's not surprising that the cantor function is easier to fit than the Weirestress function. The later contains more information in its creeks.



1.6 Conclusion

In conclusion, the smooth, low frequency, non-periodic functions are more neural network friendly. Meaning it takes less time to tain a neural network to fit them. This makes sense as they have less information encoded in them compare to their rivals. Noise increase the amount of time it takes for model parameters to converge.