CogSci131

Assignment#1, Neural Networks

February 27, 2025

1 Khan's Building in Bangla Desh and Starry Night

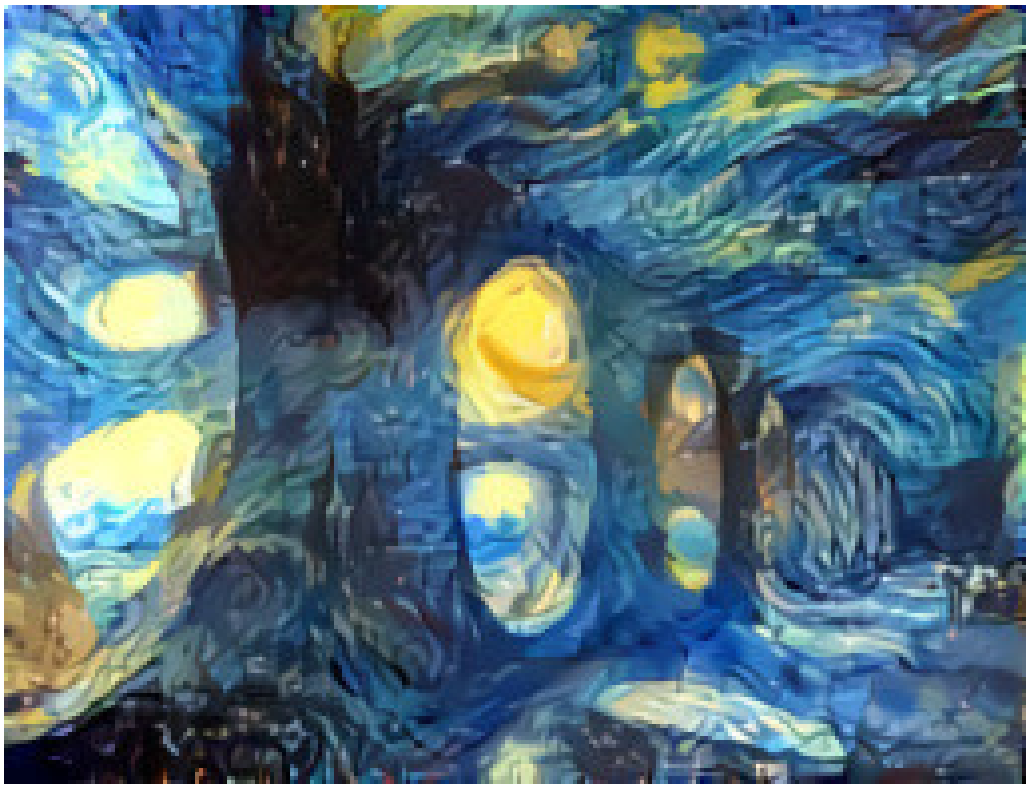

Figure 1.1: Figure Produced with Neural Net

1.1 Instructions.

Assignment #1 CogSci 131 Spring 2025

Neural Nets
Group Assignment (Teams of 2 persons)
Total number of points 200.

Instructions
This assignment will count for two coding assignment. Please submit a working jupyter notebook with the solution to the questions below. Do not use someone else's work; all code has to be your own. Use the .ipynb scripts provided in the modules section. You can submit a pdf file in addition to your ipynb file, or use the markdown language in the script.

The deadline for this assignment is Wed. March 11th at 11:59 pm . You grade will be low if you do not submit a working ipynb file! So... watch out!

1.2 First Problem. Understand the structure of a neural network.100 Points

Using the example found in neural nets demystified and in the script provided, write a program that minimizes the cost function to a given accuracy set in advance by you. Please do not get stuck in the meaning of accuracy. This simply means a given threshold that is reasonable and you think your program could reach, for example 0.1% of the target. Notice that the program does part of the job for you. In this case there are two directions:

$$\mathrm{NN}.\widehat{W_1} = \mathrm{NN}.\widehat{W_1} + scalar\frac{\partial J}{\partial \widehat{W_1}} \tag{1.1}$$

and the other direction:

$$\mathrm{NN}.\widehat{W_1} = \mathrm{NN}.\widehat{W_1} - scalar\frac{\partial J}{\partial \widehat{W_1}} \tag{1.2}$$

and the same for $\mathrm{NN}.\widehat{W_2}$. J is the cost function.

Figure 1.2: Figure Produced with Neural Net

Plot the cost vs iteration and see if you get a similar figure to the one I am showing above. Scalar is a parameter that allows to change the step of the iteration. Some people call this the learning rate for some mysterious reason. Explore what happens if the learning rate is unusually large. I provide an example in the code.

1.3 Second Problem. Use ReLu instead of sigmoid activation. 100 Points

Use the ReLu activation function instead of the sigmoid function to construct a neural net of the same dimensions used in 1.2. Compare how fast this neural network works with the one that uses the sigmoid function activation. Add a new hidden layer with 3 units and compare with the previous network. Increase the number of units in the new hidden layer to 10. Compare again. Make sure that when you use RELU, the last step of the forward propagation still uses a sigmoid activation function (This guarantees that the output of the network is still between 0 and 1, and it avoids the problem of getting stuck in negative values).