

Homework 4

601.464/664 Artificial Intelligence

Jason Zhang jzhan127

May 2, 2020

- 1 Open the following google colaboratory notebook. Follow all the steps specified in it.

Answer:

<https://colab.research.google.com/drive/1Hig8QFfLikUaYv0MauQRLDo23gZvV6Ue>

- 2 Explain the difference between fully connected layer and a convolutional layer.

Answer:

A fully connected layer has each neuron in a layer is connected to every other neuron in the next layer. A convolutional layer instead mimics a neuron receptive field instead, thus each neuron in a layer only has visibility of a few neurons of the neurons in the previous layer, specifically the ones that are within the convolutional filter(receptive field).

- 3 What is a softmax function and where is it used in neural networks?

Answer:

The softmax function is an activation function used in neural networks when it comes time to output a prediction. Mathematically softmax is defined as: $softmax(y_i) = \frac{e^{y_i}}{\sum_j e^{y_j}}$. It is used in cases of multiclass classification problems where decisions can only be one of the classes IE an image of a number can only belong to the class of one number.

- 4 Give an example of non-linearities used in neural networks. Why is it necessary to have it in networks?

Answer:

Commonly used non-linearities used in neural networks are: Tanh and Sigmoid. They are necessary due to the fact that they introduce non linearity to our neural networks. Having non-linear activation functions allow neural network allows the neural network to learn more interesting decision boundaries and thus allows for more accurate decision making. Not using these activation functions would make it so that our neural network could only learn linear decision boundaries, which clearly does not generalize well.

- 5 What are the loss functions used for regression and classification?

Answer:

For regression a commonly used loss function is Mean Squared Error. For classification a commonly used loss function is Cross Entropy Loss.

- 6 Using what algorithm gradients are usually efficiently computed in neural networks?

Answer:

The algorithm used is backpropagation with gradient descent. Essentially we use gradient descent to compute a direction towards the optimum and adjust our neural network weights based on that computed gradient. The update to these weights in a neural network is done via backpropagation.

- 7 What is the discount factor and how is it used when computing the reward in reinforcement learning?

Answer:

The discount factor γ essentially determines how much the agent will value rewards in the far

future vs in the near future. In a reinforcement learning algorithm it is used in conjunction with the rewards to help the agent evaluate its potential future gain. γ is typically between 0 and 1, where a smaller gamma corresponds to a more near sighted reward evaluation and a larger gamma corresponds to a further future reward evaluation.