Trevor Bright

Assignment 1

1/28/2020

1) $\tanh(v) = 2\,\text{sigmoid}(2v) - 1$

$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} = \frac{e^x + e^{-x} - 2e^{-x}}{e^x + e^{-x}} = 1 + \frac{-2e^{-x}}{e^x + e^{-x}}$

$= 1 + \frac{-2}{e^{2x}+1} = 1 - \frac{2}{e^{2x}+1} = 1 - 2\,\text{sigmoid}(2x) =$

$= -1(1 - 2\,\text{sigmoid}(2x)) = \boxed{2\,\text{sigmoid}(2x) - 1}$

2) a) $\text{Sigmoid}(-v) = \frac{e^{-v}}{e^{-v}+1}$ $\quad \phi(-v) = 1 - \phi(v)$

$\text{sigmoid}(v) = \frac{e^v}{e^v+1} \Rightarrow 1 - \frac{e^v}{e^v+1} \Rightarrow \frac{1+e^v}{1+e^v} - \frac{e^v}{1+e^v} = \frac{1}{1+e^v}$ $\left(\frac{e^{-v}}{e^{-v}}\right)$

$= \frac{e^{-v}}{e^{-v}+1} = \text{sigmoid}$

b) Tanh activation: $\phi(-v) = -\phi(v)$

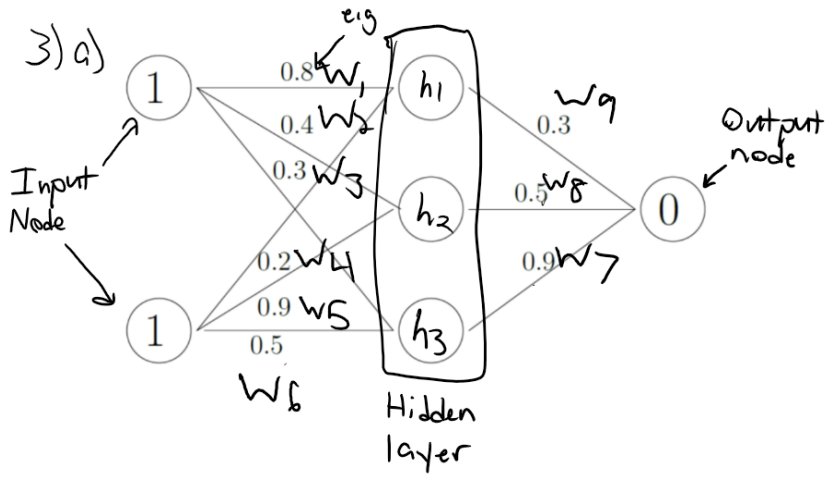
Tanh activ. $= \frac{e^x - e^{-x}}{e^x + e^{-x}}$

$\phi(-x) = \frac{e^{-x} - e^x}{e^{-x} + e^x}(-1) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$

c) Hard tanh activation: $\phi(-v) = -\phi(v)$

$\text{hardtanh}(v) = \begin{cases} -1 : v \le -1 \\ v : -1 \le v \le 1 \\ 1 : v > 1 \end{cases}$ $\quad \text{hardtanh}(-v) = \begin{cases} -1 : -v < -1 \\ -v : -1 \le -v \le 1 \\ 1 : -v > 1 \end{cases}$ (-1) (-1)

$(-1)\,\text{hardtanh}(v) = \begin{cases} 1 : v > 1 \\ v : 1 \ge v \ge -1 \\ -1 : v < -1 \end{cases}$



Input node: This is the point where input data is received.
Weights: These determine how important each of the nodes are. This is the only thing we can change in our neural network.
Hidden layer: Does the computation. It adds the weights to the input from all or a select range of activation function outputs
Output node: The point of output after processing data in hidden layer.

b) Sigmoid $= \frac{1}{1+e^{-x}}$

$X_1 = (.8)(1)+(.2)(1) = 1$ $\quad X_2 = .4(1)+.9(1) = 1.3$

$h_1 = \frac{1}{1+e^{-1}} = .731058$ $\quad h_2 = .785834983$

$X_3 = .3(1)+(.5)(1) = .8$

$h_3 = \frac{1}{1+e^{.8}} = .689974$

$X_0 = .731058(.3)+.785834(.5)$
$+.689974(.9) = 1.233212098$

$h_0 = \frac{1}{1+e^{-1.233}} = .774380$

$E = (0-.774380)^2 = .5996648$

C)

.5 seems to be a common learning rate

$$E_{w_9} = \frac{\partial E_{total}}{\partial w_9} * \frac{\partial Out_o}{\partial O} * \frac{\partial O}{\partial w_9}$$

$$\frac{\partial E_{total}}{\partial w_9} = 2(0 - Out_{h1}) * (-1 * 0)$$
$$= 2(0 - .731058) = 1.54876$$

$$\frac{\partial Out_o}{\partial O} = Out_o(1 - Out_o) = .774(1 - .774) = .174715$$

$$\frac{\partial O}{w_9} = 1 * Out_{h1} * w_9^{0} + 0 + 0$$
$$= Out_{H1} = .774380$$

$$E_{w_9} = 1.5487 * .174715 * .774380$$
$$= .209541358 7$$

$$w_9 = .3 - (.5)(.198) = .201$$

$$E_{w_8} = \frac{\partial E_T}{\partial w_8} * \frac{\partial Out_o}{\partial O} * \frac{\partial O}{\partial w_8}$$

$$\frac{\partial E_T}{\partial w_8} = 2(0 - .785834) * (-1 * 0) = 1.571668$$

$$\frac{\partial Out_o}{\partial O} = .174715$$

$$\frac{\partial O}{\partial w_8} = .786$$

$$E_{w_8} = (1.572)(.175)(.786) = .213$$

$$w_8 = .5 - (.5)(.213) = .3935$$

$$E_{w_7} = \frac{\partial E_T}{\partial w_7} * \frac{\partial Out_o}{\partial O} * \frac{\partial O}{\partial w_7} = (-2(.774))(.774(1 - .774))(.680)$$
$$= .187$$

$$w_7 = .9 - (.5)(.187) = .807$$

$$E_{w_1} = \frac{\partial E_T}{\partial w_1} = \frac{\partial E_T}{\partial out_{h_1}} \left(\frac{\partial out_{h_1}}{\partial h_1}\right)\left(\frac{\partial h_1}{\partial w_1}\right)$$

$$\frac{\partial E_1}{\partial out_{h_1}} = \frac{\partial E_1}{\partial O}\left(\frac{\partial O}{\partial out_{h_1}}\right) = .271(.03) = 0.081$$

$$\frac{\partial E_1}{\partial O} = \frac{\partial E_1}{\partial out_O}\left(\frac{\partial out_O}{\partial O}\right) = 1.548(.175) = .271$$

$$\frac{\partial E_2}{\partial O} = \frac{\partial E_2}{\partial out_O}\left(\frac{\partial out_O}{\partial O}\right) = 1.548(.175) = .271$$

$$\frac{\partial E_2}{\partial out_{h_2}} = \frac{\partial E_2}{\partial O}\left(\frac{\partial O}{\partial out_{h_2}}\right) = .271(0.5) = .136$$

$$\frac{\partial E_3}{\partial O} = \frac{\partial E_3}{\partial out_O}\left(\frac{\partial out_O}{\partial O}\right) = 1.548(.175) = .271$$

$$\frac{\partial E_3}{\partial O} = \frac{\partial E_3}{\partial O}\left(\frac{\partial O}{\partial out_{h_3}}\right) = .271(0.9) = .244$$

$$\frac{\partial E_T}{\partial out_{h_1}} = \frac{\partial E_1}{\partial out_{h_1}} + \frac{\partial E_2}{\partial out_{h_2}} + \frac{\partial E_3}{\partial out_{h_3}} = .461$$

$$\frac{\partial out_{h_1}}{\partial h_1} = out_{h_1}(1 - out_{h_1}) = .73(1 - .731) = .197$$

$$\frac{\partial h_1}{\partial w_1} = x_1 = 1$$

$$E_1 = \frac{\partial E_T}{\partial out_{h_1}}\left(\frac{\partial out_{h_1}}{\partial h_1}\right)\left(\frac{\partial h_1}{\partial w_1}\right) = .461(.197)(1) = .091 = E_4$$

$$w_1 = .8 - (.5)(.091) = .755$$

$$E_2 = \frac{\partial E_T}{\partial out_{h_2}}\left(\frac{\partial out_{h_2}}{\partial h_2}\right)\left(\frac{\partial h_2}{\partial w_2}\right) = .461(.786(1 - .786))(1) = .078 = E_5$$

$$w_2 = .4 - (.5)(0.78) = .361$$

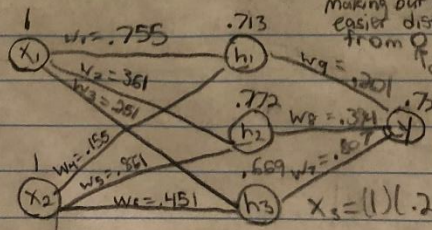
$$E_3 = \frac{\partial E_T}{\partial out_{h_3}}\left(\frac{\partial out_{h_3}}{\partial h_3}\right)\left(\frac{\partial h_3}{\partial w_3}\right) = .461(.690(1 - .690))(1) = .099 = E_6$$

$$w_3 = .3 - (.5)(.099) = .251$$

$$w_4 = .2 - (.5)(.091) = .155$$

$$w_5 = .9 - (.5)(.078) = .861$$

$$w_6 = .5 - (.5)(.099) = .451$$



making out node y for easier distinction from 0 and zero out node

$$x_1 = (1)(.755) + (1)(.155) = .91$$

$$h_1 = \frac{1}{1 + e^{-.91}} = .713$$

$$x_2 = (1)(.361) + (1)(.861) = 1.22$$

$$h_2 = \frac{1}{1 + e^{-1.22}} = .772$$

$$x_3 = (1)(.251) + (1)(.451) = .702$$

$$h_3 = \frac{1}{1 + e^{-.702}} = .669$$

$$x_y = (.713)(.201) + (.772)(.391) + (.669)(.80) = .857$$

$$y = \frac{1}{1 + e^{-.857}} = .728$$

$$E = (0 - .728)^2 = .530$$

Pass 2 $\frac{\partial T}{\partial out_y} = -2(0-.728) = 1.456$

$\frac{\partial out_y}{\partial y} = .728(1-.728) = .198$

$E_7 = 1.456(.198)(.669) = .193$

$w_7 = .807 - (.5)(.193) = .711$

$E_8 = 1.456(.198)(.772) = .223$

$w_8 = .394 - (.5)(.223) = .283$

$E_9 = 1.456(.198)(.713) = .206$

$w_9 = .201 - (0.5)(.206) = .098$

$\frac{\partial T}{\partial out_{h1}} = \frac{\partial E_1}{\partial out_{h1}} + \frac{\partial E_2}{\partial out_{h2}} + \frac{\partial E_3}{\partial out_{h3}} = 0.058 + 0.113 + 0.232 = 0.403$

$\frac{\partial E_1}{\partial y} = 1.456(.198) = .288$ $\frac{\partial E_1}{\partial out_{h1}} = .288(.201) = .058$

$\frac{\partial E_2}{\partial y} = 1.456(.198) = .288$ $\frac{\partial E_2}{\partial out_{h2}} = .288(.394) = .113$

$\frac{\partial E_3}{\partial y} = .1456(.198) = .288$ $\frac{\partial E_3}{\partial out_{h3}} = .288(.807) = .232$

$E_1 = (.461)(.713(1-.713))(1) = .094 = E_4$

$w_1 = .755 - (.5)(.094) = .708$

$E_2 = (.461)(.772(1-.772))(1) = .081 = E_5$

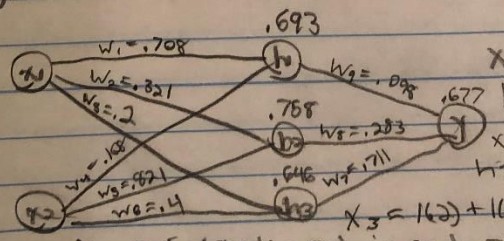
$w_2 = .361 - 0.5(.081) = .321$

$E_3 = .461(.669(1-.669))(1) = .102 = E_6$

$w_3 = .251 - 0.5(.102) = .2$

$w_4 = .155 - (0.5)(.094) = .108$

$w_5 = .861 - .5(.081) = .821$ $w_6 = .451 - .5(.102) = .4$



$x_1 = (1)(.708) + 1(.108) = .816$ $h_1 = \frac{1}{1+e^{-.816}} = .693$

$x_2 = 1(.321) + 1(.821) = 1.142$ $h_2 = \frac{1}{1+e^{-1.142}} = .758$

$x_3 = 1(.2) + 1(.4) = .6$ $h_3 = \frac{1}{1+e^{.6}} = .646$

$x_y = .693(.098) + .758(.283) + (.646)(.711) = .742$

$y = \frac{1}{1+e^{-.742}} = .677$

$E = (0-.677)^2 = .458$ or 45.8%

4. a)

```
(tf_gpu) PS C:\Users\Trevor\OneDrive\Spring 2020\Deep Learning\HW1> python .\HW1-4a.py
2020-01-28 18:02:47.745986: I tensorflow/stream_executor/platform/default/dso_loader.cc:44] Successfully opened dynamic
library cudart64_100.dll
WARNING:tensorflow:From C:\Users\Trevor\Anaconda3\envs\tf_gpu\lib\site-packages\tensorflow_core\python\keras\initializer
s.py:143: calling RandomNormal.__init__ (from tensorflow.python.ops.init_ops) with dtype is deprecated and will be remov
ed in a future version.
Instructions for updating:
Call initializer instance with the dtype argument instead of passing it to the constructor
WARNING:tensorflow:From C:\Users\Trevor\Anaconda3\envs\tf_gpu\lib\site-packages\tensorflow_core\python\ops\resource_vari
able_ops.py:1630: calling BaseResourceVariable.__init__ (from tensorflow.python.ops.resource_variable_ops) with constrai
nt is deprecated and will be removed in a future version.
Instructions for updating:
If using Keras pass *_constraint arguments to layers.
WARNING:tensorflow:From C:\Users\Trevor\Anaconda3\envs\tf_gpu\lib\site-packages\tensorflow_core\python\ops\nn_impl.py:18
3: where (from tensorflow.python.ops.array_ops) is deprecated and will be removed in a future version.
Instructions for updating:
Use tf.where in 2.0, which has the same broadcast rule as np.where
2020-01-28 18:03:03.978817: W tensorflow/stream_executor/platform/default/dso_loader.cc:55] Could not load dynamic libra
ry 'nvcuda.dll'; dlerror: nvcuda.dll not found
2020-01-28 18:03:04.006484: E tensorflow/stream_executor/cuda/cuda_driver.cc:318] failed call to cuInit: UNKNOWN ERROR (
303)
2020-01-28 18:03:04.035995: I tensorflow/stream_executor/cuda/cuda_diagnostics.cc:169] retrieving CUDA diagnostic inform
ation for host: DESKTOP-DK5D7US
2020-01-28 18:03:04.057408: I tensorflow/stream_executor/cuda/cuda_diagnostics.cc:176] hostname: DESKTOP-DK5D7US
2020-01-28 18:03:04.082302: I tensorflow/core/platform/cpu_feature_guard.cc:142] Your CPU supports instructions that thi
s TensorFlow binary was not compiled to use: AVX AVX2
Standardized: 97.50% (4.03%)
(tf_gpu) PS C:\Users\Trevor\OneDrive\Spring 2020\Deep Learning\HW1>
```

b)

```
(tf_gpu) PS C:\Users\Trevor\OneDrive\Spring 2020\Deep Learning\HW1> python .\HW1-4b.py
2020-01-28 18:06:11.822870: I tensorflow/stream_executor/platform/default/dso_loader.cc:44] Successfully opened dynamic
library cudart64_100.dll
WARNING:tensorflow:From C:\Users\Trevor\Anaconda3\envs\tf_gpu\lib\site-packages\tensorflow_core\python\keras\initializer
s.py:143: calling RandomNormal.__init__ (from tensorflow.python.ops.init_ops) with dtype is deprecated and will be remov
ed in a future version.
Instructions for updating:
Call initializer instance with the dtype argument instead of passing it to the constructor
WARNING:tensorflow:From C:\Users\Trevor\Anaconda3\envs\tf_gpu\lib\site-packages\tensorflow_core\python\ops\resource_vari
able_ops.py:1630: calling BaseResourceVariable.__init__ (from tensorflow.python.ops.resource_variable_ops) with constrai
nt is deprecated and will be removed in a future version.
Instructions for updating:
If using Keras pass *_constraint arguments to layers.
WARNING:tensorflow:From C:\Users\Trevor\Anaconda3\envs\tf_gpu\lib\site-packages\tensorflow_core\python\ops\nn_impl.py:18
3: where (from tensorflow.python.ops.array_ops) is deprecated and will be removed in a future version.
Instructions for updating:
Use tf.where in 2.0, which has the same broadcast rule as np.where
2020-01-28 18:06:17.308231: W tensorflow/stream_executor/platform/default/dso_loader.cc:55] Could not load dynamic libra
ry 'nvcuda.dll'; dlerror: nvcuda.dll not found
2020-01-28 18:06:17.321935: E tensorflow/stream_executor/cuda/cuda_driver.cc:318] failed call to cuInit: UNKNOWN ERROR (
303)
2020-01-28 18:06:17.345021: I tensorflow/stream_executor/cuda/cuda_diagnostics.cc:169] retrieving CUDA diagnostic inform
ation for host: DESKTOP-DK5D7US
2020-01-28 18:06:17.358441: I tensorflow/stream_executor/cuda/cuda_diagnostics.cc:176] hostname: DESKTOP-DK5D7US
2020-01-28 18:06:17.367711: I tensorflow/core/platform/cpu_feature_guard.cc:142] Your CPU supports instructions that thi
s TensorFlow binary was not compiled to use: AVX AVX2
Standardized: 99.00% (2.00%)
(tf_gpu) PS C:\Users\Trevor\OneDrive\Spring 2020\Deep Learning\HW1>
```