**AI51801/CSE54501 Deep generative model/Advanced computer vision Quiz 1**

**Link to this document (shared in the Blackboard/announcement):**

**https://docs.google.com/document/d/1BydS939JWDDHWZaVEKQsCvskXaJBHOqZT_eGvkC7me8/edit?usp=sharing**

**Duration: 3/22 (Tue) 4:50PM-6:00PM**

Answer sheet link: https://forms.gle/ggSnZ7p1KsxQ871K9
(If you need to update your answer, you can modify the original answers or can re-submit by filling only new answers. We will reflect the most recent answer for each problem.)

1. **(40pts) Please indicate 'True' or 'False' for the sentences below.**

**[Scores are given as follows: no answer (0 point), wrong answer (-6 points), correct answer (+10 points)]**

a. (10pts) Outputs of the 2D convolutional layers are differentiable with respect to their inputs. **(True)**

b. (10pts) For the optimizer, optimizer.step() (→ loss.backward() ) is a function which is called for calculating gradients and assigning gradients to .grad instance of pytorch variables. **(False)**

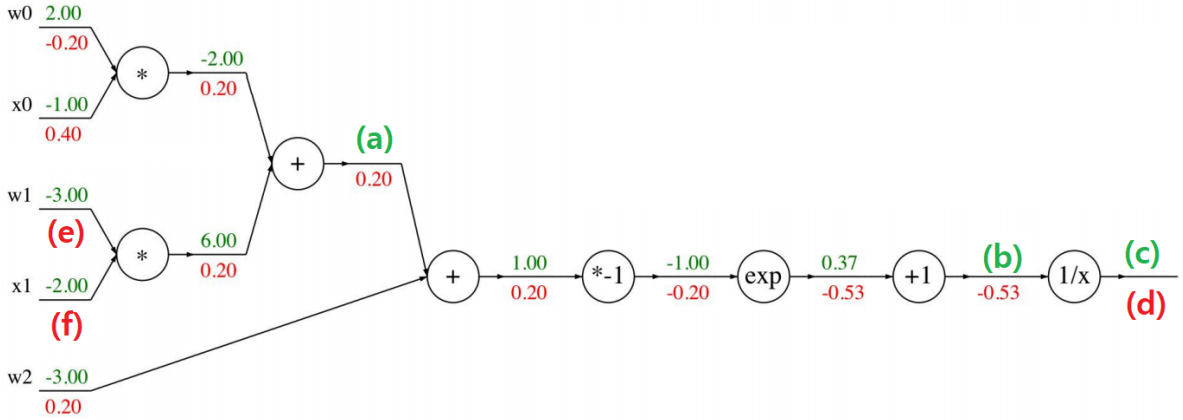
c. (10pts) Before calling loss.backward(), gradients need to be initialized to NULL and the function called for that purpose is: optimizer.init()(→ optimizer.zero_grad() ). **(False)**

d. (10pts) Stochastic gradient descent goes many non-optimal steps; however it converges better than gradient descent in practice. **(True)**

**2. (60pts)**

$$f(w,x) = \frac{1}{1+e^{-(w_0x_0+w_1x_1+w_2)}}$$

**Below is the graph built for calculating the forward and backward pass of the above function f(w,x). Green and red fonts are indicating forwarded values and gradients, respectively. Given that w0, x0, w1, x1 and w2 are initialized to 2.0, -1.0, -3.0, -2.0 and -3.0, please fill (a)-(f) with proper numerical values. (each 10pts.) Calculate until 2 digits are obtained below the floating point.**




(a) 4 (sol: -2.0+6.0=4.0)
(b) 1.37 (sol: 0.37+1=1.37)
(c) 0.73 (sol: 1/1.37=0.73)
(d) 1 (sol: this is always 1.00, as it is the initial node)
(e) -0.4 (sol: if f=w1*x1=-2.0*x1, the gradient for w1 is 0.2 * df/dw1 = 0.2*-2 = -0.4)
(f) -0.6 (sol: if f=w1*x1=w1*-3.0, the gradient for x1 is 0.2 * df/dx1 = 0.2*-3 = -0.6)

**3. (No point.) Please share your opinion: Do you agree with changing this course into an offline course, after the mid-term? Please select `yes' or `no' and provide some reasons (You can use Korean to share your opinion).**