CSC 8851 Deep Learning

Assignment 1

Due Date: 9/3/2021 (by 11:59 pm)

Name: _____   Campus ID: _____

Background Knowledge Test

Problem 1. (1 point) We are machine learners with a slight gambling problem (very different from gamblers with a machine learning problem!). Our friend, Bob, is proposing the following payout on the roll of a dice:

$$payout = \begin{cases} \$1 & x = 1 \\ -\$1/4 & x \neq 1 \end{cases}$$

Where $x \in \{1,2,3,4,5,6\}$ is the outcome of the roll, (+) means payout to us and (-) means payout to Bob. Is this a good bet for us? Are we expected to make money?

Problem 2. (1 point) X is a continuous random variable with the probability density function

$$p(x) = \begin{cases} 4x & when\ 0 \leq x \leq 1/2 \\ -4x + 4 & when\ 1/2 \leq x \leq 1 \end{cases}$$

What is the equation for the corresponding cumulative density function (cdf) $C(x)$?
[Hint: Recall that CDF is defined as $C(x) = P(X \leq x)$.]

Problem 3. (1 point) Recall that the variance of a random variable is defined as $Var[X] = E[(X - \mu)^2]$, where $\mu = E[X]$. Use the properties of expectation to show that we can rewrite the variance of a random variable X as

$$Var[X] = E[X^2] - (E[X])^2$$

Problem 4. (1 point) A random variable x in standard Gaussian distribution has following probability density

$$p(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

Evaluate following integral

$$\int_{-\infty}^{\infty} p(x)(ax^2 + bx + c)dx$$

[Hint: This is not a calculus question!]

Problem 5. (3 points) Consider the following function of $x = (x_1, x_2, x_3, x_4, x_5, x_6)$:

$$f(x) = \sigma\left(\log\left(5\left(\max\{x_1, x_2\}.\frac{x_3}{x_4} - (x_5 + x_6)\right)\right) + \frac{1}{2}\right)$$

where log(x) is the natural log function, and σ is the sigmoid function

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

Evaluate $f(x)$ at $\hat{x} = (5, -1, 6, 12, 7, -5)$. Then, compute the gradient $\nabla_x f(x)$ and evaluate it at the same point.

Problem 6. (3 points) Set up your Python programming environment by following the instruction below.

Setting up a Python programming environment for deep learning can be challenging. Depending on the system and privilege you have, it can be as easy as running a few scripts or completely a headache. In this class, we will use Python3+ and Numpy for programming assignments. In generally, you can set up your environment in Linux, Mac and Windows. Linux and Mac are preferred as they have better support to python and DL frameworks; you may just need to run a few scripts and complete this task. Otherwise, you may set up your environment in Windows, but it's a little bit more involved. So the windows' route is not recommended, and you should avoid it whenever possible.

0.   Download assignment1.zip from iCollege "homework" section

Linux (Preferred. Setup in Mac is similar to this.):
0.   set up python3 environment (e.g., <u>install virtual environment</u> or <u>install anaconda</u>)
1.   unzip assignment1.zip
2.   cd assignment1
3.   pip install -r requirements.txt (if you don't have sudo, you can install virtual box or anaconda)
4.   cd assignment1/lib/datasets
5.   source get_datasets.sh
6.   cd assignment1
7.   jupyter notebook (and save your results)
8.   If your notebook runs without problem, then press "ctrl + s" to save your notebook, run ". collectSubmission.sh" to collect your results, and send your zip file along with your completed problem set to iCollege hw1 dropbox.

Windows: (Not recommended, but just for your information)
•   Install Cygwin https://cygwin.com/install.html

Select packages: wget, zip and unzip

Cygwin Setup - Select Packages

Select Packages
Select packages to install

View: Category ▼   Search: wget   Clear

Category	Current	New	Bin?	Src?	Size	Package
⊟ All ⟳ Default						
⊞ Debug ⟳ Default						
⊞ Perl ⟳ Default						
⊞ Utils ⟳ Default						
⊟ Web ⟳ Default						
	1.19.1-2	⟳ 1.19-1	☒	☐	794k	wget: Utility to retrieve files from the WWW via HTTP and FTP

Cygwin Setup - Select Packages

Select Packages
Select packages to install

View: Category ▼   Search: zip   Clear

Category	Current	New	Bin?	Src?	Size	Package
⊟ All ⟳ Default						
⊟ Archive ⟳ Default						
	1.0.6-3	⟳ Keep	n/a	☐	35k	bzip2: BZip file de/compressor
		⟳ Skip	n/a	n/a	15k	liblrzip-devel: Long Range ZIP archive library (development)
		⟳ Skip	n/a	n/a	98k	liblrzip0: Long Range ZIP archive library (runtime)
		⟳ Skip	n/a	n/a	69k	libzzip-devel: ZIP file library
		⟳ Skip	n/a	n/a	20k	libzzip0.13: ZIP file library
		⟳ Skip	n/a	n/a	174k	lrzip: Long Range ZIP archiver
		⟳ Skip	n/a	n/a	77k	lzip: Lossless data compressor based on the LZMA algorithm
		⟳ Skip	n/a	n/a	72k	lziprecover: Lossless data compressor based on the LZMA algorithm
		⟳ Skip	n/a	n/a	965k	p7zip: A file archiver with very high compression ratios
		⟳ Skip	n/a	n/a	37k	pbzip2: Parallel BZIP2 de/compressor
		⟳ 6.0-17	☒	☐	183k	unzip: Info-ZIP decompression utility
		⟳ 3.0-12	☒	☐	217k	zip: Info-ZIP compression utility
		⟳ Skip	n/a	n/a	51k	zziplib: ZIP file utilities
⊞ Base ⟳ Default						
⊞ Debug ⟳ Default						
⊞ Devel ⟳ Default						
⊞ Doc ⟳ Default						
⊞ Libs ⟳ Default						
⊞ Perl ⟳ Default						
⊟ PHP ⟳ Default						
		⟳ Skip	n/a	n/a	46k	php-zip: PHP zip extension
⊟ Utils ⟳ Default						
		⟳ Skip	n/a	n/a	86k	fcrackzip: Password cracker for zip archives
		⟳ Skip	n/a	n/a	53k	rzip: Compression program to use long distance redundancies in files

1. Run cygwin64 Terminal
2. Unzip assignment1.zip

- Install Anaconda https://www.anaconda.com/download/
1. Run "**Anaconda Powershell Prompt**" as Adminstrator
2. cd assignment1
3. pip install -r requirements.txt
4. cd assignment1/lib/datasets
5. source get_datasets.sh
6. cd assignment1
7. jupyter notebook (and save your results)
8. If your notebook runs without problem, press "ctrl + s" to save your notebook, run ".collectSubmission.sh" to collect your results and send your zip file along with your completed problem set to iCollege hw1 dropbox.