Problem Set 1

CS 4277: Deep Learning

Name (print clearly): _____ Section: (e.g., 01) _____

Signature: _____

Student account username (e.g., msmith3): _____

Signing signifies that you agree to comply with the **Academic Honor Code** and course policies stated in the syllabus.

Choose one of these two options for turn-in:

1. Print this document, write or answers, scan your finished homework to a PDF, name the PDF `cs4277-ps1-<your-student-account-username>.pdf`, e.g., `cs4277-ps1-msmith3.pdf` and submit the PDF to the assignment on D2L.

2. While viewing this document in your web browser, in the address bar change `.pdf` to `.tex`, save the LaTeX source as a text file, add your answers in appropriate LaTeX markup in the appropriate spaces, compile to a PDF named as in the instructions above, and submit the PDF file to the assignment on D2L. (Whether you choose this option or not, LOOK AT THE LaTeX SOURCE!)

Question:	1	2	3	4	5	6	Total
Points:	40	15	15	15	15	15	115
Score:							

1. According to Tom Mitchell, machine learning is the study of algorithms that

   - improve their performance `P`
   - at some task `T`
   - with experience `E`.

   A well-defined learning task is given by `<P, T, E>`.

   Formulate the following problems according to Tom Mitchell's machine learning problem specification (see Machine Learning Slides) and the specification our textbook. For each of the following problems specify:

   - The task `T`,
   - The performance measure `P`,
   - The experience `E`,
   - The target function $f : \mathcal{X} \to \mathcal{Y}$, that is,
     - the input space $\mathcal{X}$, and
     - the output space $\mathcal{Y}$.

   Remember that a function maps a domain to a co-domain, and these domains are sets.

   Note: there are many correct solutions.

(a) (10 points) Medical diagnosis: A patient walks in with a medical history and some symptoms, and you want to identify the problem.

(b) (10 points) Handwritten digit recognition (for example postal zip code recognition for mail sorting).

(c) (10 points) Determining if an email is spam or not.

(d) (10 points) Predicting how an electric load varies with price, temperature, and day of week.

2. (15 points) (Problem 2.1 from our textbook) To walk "downhill" on the loss function (equation 2.5 – reproduced here for convenience),

$$L(\phi) = \sum_{i=1}^{I} (\phi_0 + \phi_1 x_i - y_i)^2 \tag{1}$$

we measure its gradient with respect to the parameters ϕ_0 and ϕ_1. Calculate expressions for the slopes $\frac{\partial L}{\partial \phi_0}$ and $\frac{\partial L}{\partial \phi_1}$. Show your work for partial credit.

Tips:

- See Partial Derivatives and Gradients in Dive Into Deep Learning for a refresher on the essential vector calculus needed for this question.

3. (15 points) (Problem 3.1 from our textbook) What kind of mapping from input to output would be created if the activation function in Equation 3.1 were linear so that $a(z) = \psi_0 + \psi_1 z$? What kind of mapping would be created if the activation function was removed, so $a(z) = z$?

4. (15 points) (Problem 3.2 from our textbook) For each of the four linear regions in Figure 3.3j, indicate which hidden units are inactive and which are active (i.e., which do and do not clip their inputs).

5. (15 points) (Problem 3.12 from our textbook) How many parameters does the model in Figure 3.7 have?

6. (15 points) (Problem 3.16 from our textbook) Write out the equations for a network with two inputs, four hidden units, and three outputs. Draw this model in the style of Figure 3.11.