COMP 433 Introduction to Deep Learning (Fall 2025)

Major Assignment (Individual) #1

Due: 11:59PM, September 28th, 2025

Note Your will be submitting two separate files from this assignment as follows:

(a) **One(1) .pdf file:** containing answers to questions as well as reported results from coding you develop. Therefore, include snapshots of the output of the code you developed. **Please write all steps**. It is not necessary to type your answers. You may write them by hand in a legible manner and include photos of your solution as part of your submission.

(b) **One(1) .zip folder:** containing all developed Python codes. As mentioned, the snapshots of the outputs should be included in the PDF file.

Theoretical Questions

Question 1 You have been hired as a data scientist at ABC Bank to apply Artificial Intelligence (AI) for improving the Bank's operations. The bank has provided you with a dataset containing information on 100,000 clients, including various features related to their profiles and activities.

Upon reviewing the dataset, you notice that some columns (e.g., name, full address, date of birth, etc) are encrypted.

(a) Why do you think these columns have been encrypted? What impact could this have on your machine learning tasks?

(b) Propose an AI task (model) that could be implemented using traditional programming (e.g., expert systems). Explain why this task does not necessarily require machine learning.

(c) Suggest a beneficial supervised **regression** task that could be performed with this dataset.

(d) Suggest a beneficial supervised **classification** task that could be performed with this dataset.

(e) Suggest a beneficial **unsupervised** learning task that could be performed with this dataset.

(f) For the supervised tasks you proposed, what would be a typical split between training, validation, and test sets? Briefly explain the purpose of each set.

(g) Propose an AI task for which deep learning would be more suitable than traditional machine learning. Explain why deep learning may be necessary in this case.

Question 2 Provide a comprehensive answer to each of the following questions.

(a) In a neural network, why do we use activation functions? In other words, why don't we simply pass the weighted sum of the inputs directly to the output?

(b) Compare the ReLU activation function with the Leaky ReLU activation function.

Question 3 Consider the following dataset with three (3) samples and two (2) features. We are performing a linear regression task using gradient descent

$$\hat{y} = w_1 x_1 + w_2 x_2 + w_0$$

After a few iterations, suppose the model parameters are currently as follows:

$$w_1 = 2, \quad w_2 = -1, \quad w_0 = 3$$

Sample	x_1	x_2	True Label (y)
1	1	2	4
2	2	1	3
3	0	3	1

(a) Compute the Mean Squared Error (MSE).

(b) Assume that the weights are updated after processing each data point (i.e., with three samples, three updates in each epoch). Compute the the updated weights after each sample in the next epoch, assuming the learning rate $\eta = 0.1$. (note: for a given data point, a weight is updated as $w_i \leftarrow w_i - \eta \frac{\partial MSE}{\partial w_i}$. The gradients for this case can be found as follows: $\frac{\partial MSE}{\partial w_i} = (\hat{y} - y)x_i$ and $\frac{\partial MSE}{\partial w_0} = (\hat{y} - y))$

Question 4 Consider a sentiment analysis model that classifies sentences into three categories: Positive, Negative, or Neutral (three classes). For a dataset with five sentences, the neural network outputs the following probability distributions over the three classes.

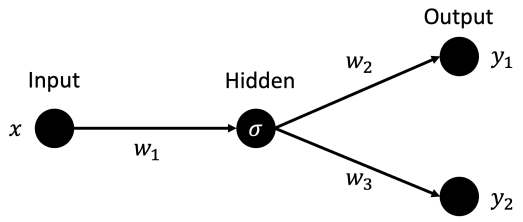
Sample	**Positive**	**Negative**	**Neutral**	**True Label**
1	0.70	0.20	0.10	Positive
2	0.10	0.80	0.10	Negative
3	0.30	0.40	0.30	Neutral
4	0.60	0.25	0.15	Positive
5	0.20	0.30	0.50	Neutral

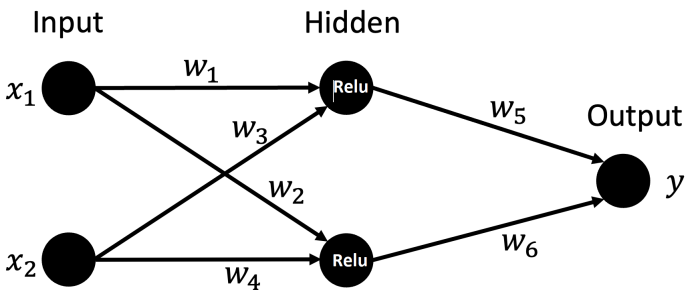
Table 1: Model predicted probabilities for sentiment analysis (3 classes) with true labels.

(a) Using categorical cross-entropy with natural logarithm, calculate the **average** loss for this batch of five samples. Write all the steps.

(b) Assuming the model chooses the class of the highest probability as the output class, what is the classification accuracy of the classification model represented in the table?

Question 5 For each of the following neural networks, derive an expression for the output of the neural network in terms of the input and the weights.





where σ is the sigmoid activation function and Relu indicates the Relu activation function.

Implementation Questions

Question 1 Using the provided modified dataset of Titanic passengers (original dataset available here), build a machine learning model to predict the survival outcome. Based on information about each passenger—such as sex, age, ticket class, and fare—your model should determine whether the passenger survived or not. Note: Use the random seed (random state) of 433.

(a) Use statistical methods and graphs/plots to describe your daataset.

(b) Split your dataset into train and test sets with a 80:20 ratio. Use the train_test_split tool from scikit-learn.

(c) Build and train a Logistic Regression model using scikit-learn. Explore the parameters of the model in scikit-learn, and aim for higher classification accuracies.

(d) Report and discuss the performance of your developed model on both the train and test sets (separately). You can use scikit-learn's classification report tool.

Question 2 We use PyTorch to build deep neural networks; however, scikit-learn also provides a very simple implementation of a single neuron (perceptron). Repeat the previous question using the scikit-learn Perceptron classifier.

(a) Using the same train and test split, build and train a Perceptron classifier model using scikit-learn. Explore the parameters of the model in scikit-learn, and aim for higher classification accuracies.

(b) Report and discuss the performance of your developed model on both the train and test sets (separately). You can use scikit-learn's classification report tool.