[ENXXXX] — [Module Descriptive Name]

Student Name: [ABC DEFGH]   **Submitted Date:** [April 02, 2021]
Student Number: [123456X]   **Assignment Number:** [1]

The necessary changes to be executed, in order to transform this template to your assignment could be found within square brackets - [ ] or with the initiation of the word "Sample". The compact style of documentation is ideal to follow in an assignment where the page limit is defined and critical.

Sample Text: Entire code flow for the assignment is accessible via,
`[TheOpen-sourcelinkfortheassignmentifany]`

Problem 1   Sample Question: A part of the code for a linear classifier for CIFAR10 given in listing 1. For our linear classifier, the score function is $f(x) = Wx + b$, and the loss function is the mean sum of squared errors function.

Sample Answer: Lorem Ipsum is simply dummy text of the printing and typesetting industry. Lorem Ipsum has been the industry's standard dummy text ever since the 1500s, when an unknown printer took a galley of type and scrambled it to make a type specimen book. It has survived not only five centuries, but also the leap into electronic typesetting, remaining essentially unchanged. It was popularised in the 1960s with the release of Letraset sheets containing Lorem Ipsum passages, and more recently with desktop publishing software like Aldus PageMaker including versions of Lorem Ipsum. [1] (The citations could be added using cite package).

(a) Sample Sub-question: Implement gradient descent and run for 300 epochs.

Sample Answer: Lorem Ipsum is simply dummy text of the printing and typesetting industry. Lorem Ipsum has been the industry's standard dummy text ever since the 1500s, when an unknown printer took a galley of type and scrambled it to make a type specimen book. It has survived not only five centuries, but also the leap into electronic typesetting, remaining essentially unchanged.

The mathematical equations could be included using math package as indicated here.

$$f(x_i, W, b) = Wx_i + b \tag{1}$$

Mathematical expressions could be added within the text passage as follows. Sample Text: Here, $x_i \in R^D$ each associated with a label y_i. Here $i = 1...N$ and $y_i \in 1...K$ such that N examples, with dimensionality D, in accordance with K distinct categories.

Lorem Ipsum is simply dummy text of the printing and typesetting industry. Lorem Ipsum has been the industry's standard dummy text ever since the 1500s, when an unknown printer took a galley of type and scrambled it to make a type specimen book. It has survived not only five centuries, but also the leap into electronic typesetting, remaining essentially unchanged.

(b) Sample Sub-question: Show the weights matrix W as 10 images and (c) Report the (initial) learning rate, training and testing loss and accuracies.

Sample Answer: Lorem Ipsum is simply dummy text of the printing and typesetting

industry. Lorem Ipsum has been the industry's standard dummy text ever since the 1500s, when an unknown printer took a galley of type and scrambled it to make a type specimen book. It has survived not only five centuries, but also the leap into electronic typesetting, remaining essentially unchanged.

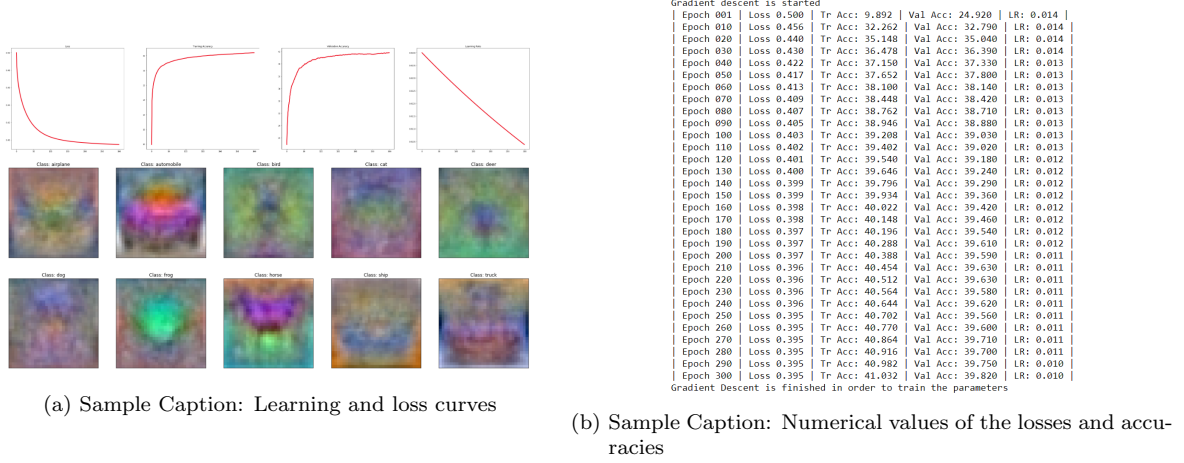*Figures could be added as follows (with sub-figures).*

(a) Sample Caption: Learning and loss curves

(b) Sample Caption: Numerical values of the losses and accuracies

Figure 1: Sample caption: Acquired results through the implemented linear classifier

Problem 2   Sample Question: Code a two-layer fully connected network with $H = 200$ hidden nodes. Choose the sigmoid function as the activation function for the hidden nodes. The output layer has no activation function.

(a) Sample Sub-question: Implement gradient descent and run for 300 epochs.
Sample Answer: **Lorem Ipsum** is simply dummy text of the printing and typesetting industry. Lorem Ipsum has been the industry's standard dummy text ever since the 1500s, when an unknown printer took a galley of type and scrambled it to make a type specimen book. It has survived not only five centuries, but also the leap into electronic typesetting, remaining essentially unchanged.

$$\frac{dl}{dw2} = \frac{dl}{dpredict} X \frac{dpredict}{dw2} \tag{2}$$

$$\frac{dl}{dw1} = \left(\frac{dl}{dpredict} X \frac{dpredict}{dh}\right) X \left(\frac{dh}{dw1x} X \frac{dw1x}{dw1}\right) \tag{3}$$

(b) Sample Sub-question: Report the (initial) learning rate, training and testing loss and accuracies.
Sample Answer: Lorem Ipsum is simply dummy text of the printing and typesetting industry. Lorem Ipsum has been the industry's standard dummy text ever since the 1500s, when an unknown printer took a galley of type and scrambled it to make a type specimen book. It has survived not only five centuries, but also the leap into electronic typesetting, remaining essentially unchanged.

Problem 3 Sample Question: Modify the code in item 2 to carry out stochastic gradient descent with a batch size of 500 (Report training and testing loss and accuracies and Compare results with item 2).

Sample Answer: Lorem Ipsum is simply dummy text of the printing and typesetting industry. Lorem Ipsum has been the industry's standard dummy text ever since the 1500s, when an unknown printer took a galley of type and scrambled it to make a type specimen book. It has survived not only five centuries, but also the leap into electronic typesetting, remaining essentially unchanged.
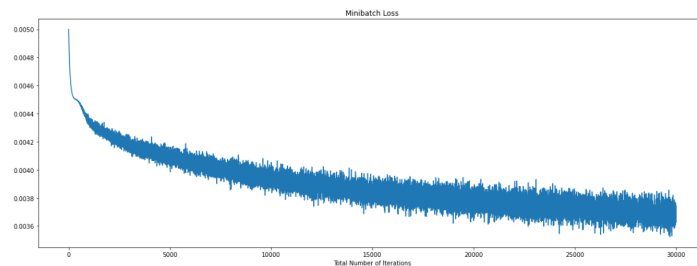

Figure 2: Sample Caption: Mini-batch loss vs total number of iterations

Lorem Ipsum is simply dummy text of the printing and typesetting industry. Lorem Ipsum has been the industry's standard dummy text ever since the 1500s, when an unknown printer took a galley of type and scrambled it to make a type specimen book. It has survived not only five centuries, but also the leap into electronic typesetting, remaining essentially unchanged.

Problem 4 Sample Question: Construct a CNN using *Keras.models.Sequential* (with the following configuration: C32, C64, C64, F64, F10. All three convolutions layers are 3x3. Max pooling (2x2) follows each convolution layer. Use SDG (with momentum) with a batch size of 50 and *CategoricalCrossentropy* as the loss (Include how many learnable parameters are there in this network?, report the parameters such as the learning rate and momentum and Report training and testing loss and accuracies)

Sample Answer: Lorem Ipsum is simply dummy text of the printing and typesetting industry. Lorem Ipsum has been the industry's standard dummy text ever since the 1500s, when an unknown printer took a galley of type and scrambled it to make a type specimen book. It has survived not only five centuries, but also the leap into electronic typesetting, remaining essentially unchanged. $loss = -\sum_{i=1}^{outputSize} y_i X \log \hat{y_i}$.

References

[1] A. Krizhevsky, "Learning multiple layers of features from tiny images," 2009. [Online]. Available: https://www.cs.toronto.edu/~kriz/learning-features-2009-TR.pdf