

2025 Spring AMA564 Assignment 2

The Hong Kong Polytechnic University

Due 23:59, Sunday, April 06, 2025

1. Let x be a input image

$$x = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \mathbf{1} & 0 & 0 & \mathbf{1} & 0 \\ 0 & \mathbf{1} & 0 & 0 & \mathbf{1} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Let f_1, f_2, f_3 be three convolutional filters

$$f_1 = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \quad f_2 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \quad f_3 = \begin{bmatrix} 1 & -1 \\ -1 & 0 \end{bmatrix}.$$

Let *MaxPool* be a max-polling operator with filter size 2×2 and stride 2. Let *AvgPool* be an average-pooling operator with filter size 3×3 and stride 3.

Question:

- (1) (1 mark) Please calculate the results of the convolution on x using filters f_1, f_2, f_3 with no padding and stride 1.
 - (2) (1 mark) Please calculate the result of the max-pooling on x using *MaxPool*.
 - (3) (1 mark) Please calculate the result of the average-pooling on x using *AvgPool*.
2. **Background:** In practice, we need to train the built deep neural network by an appropriate optimization algorithm. In this course, we introduced several popular optimization algorithms for training the deep neural network, such as stochastic gradient descent (SGD), stochastic gradient descent with momentum, ADAM, SVRG. To have a better understanding of the role of optimization algorithms in deep learning, we compare their performance in this question.

Particularly, we test their performance by training the 18-layer ResNet on the CIFAR-10 dataset. A notebook entitled “ResNet.ipynb” can be found in the attachment, which contains the code for building the 18-layer ResNet .

Question: Finish the rest parts of the codes for training the built neural network on CIFAR-10 dataset and plot the figures of training loss, test loss, training accuracy, test accuracy for the first 30 epochs.

- (1) (1 mark) Train the ResNet by SGD with momentum (with momentum parameter 0.9, learning rate = 0.001) and plot the required figures. Please include your plotted figures in your assignment solution.

- (2) **(1 mark)** Train the ResNet by ADAM (with learning rate = 0.01 and default values for β_1 and β_2) and plot the required figures. Please include your plotted figures in your assignment solution.
3. **(2 marks)** Please give answers in “True” or “False” for the following statements:
- (a) VAEs are a type of supervised learning model that can learn to map input features to output targets.
 - (b) The discriminator network in GANs is trained to distinguish between the generated data and the training data.
 - (c) RNNs can be used for time series forecasting, where they can learn to predict future values based on past observations.
 - (d) The basic building block of an RNN is a single neuron that takes a sequence of inputs and produces a single output.
 - (e) Long Short-Term Memory (LSTM) networks are a type of RNN that can retain information over a long period of time and avoid the vanishing gradient problem.
4. **(3 marks)** Multiple Choice Questions. Please choose the appropriate letters **A**, **B**, **C** or **D**.
- (a) Which of the following is a common application of generative models such as VAEs and GANs?
 - A.** Image recognition
 - B.** Speech recognition
 - C.** Text generation
 - D.** Image generation
 - (b) Which of the following is a characteristic of generative models?
 - A.** They can only be used for supervised learning tasks
 - B.** They can generate new data that is similar to the training data
 - C.** They can only be used for classification tasks
 - D.** They require a large amount of labeled data for training
 - (c) Which of the following is a type of generative model that uses two neural networks to generate new data?
 - A.** Variational autoencoders
 - B.** Generative Adversarial Networks
 - C.** Flow-based generative models
 - D.** Score-based generative model
 - (d) Which of the following is a type of RNN that can handle long-term dependencies?
 - A.** Elman recurrent neural network
 - B.** Jordan recurrent neural network
 - C.** One-to-many RNN
 - D.** Long Short-Term Memory (LSTM)
 - (e) Which of the following is a type of word embedding that is trained by predicting the context words given a target word?
 - A.** CBOW

- B. Skip-gram
- C. One-hot embedding
- D. Word2Vec