Advanced Machine Learning 2024

Final Exam

Chennai Mathematical Institute

26 November 2024, 09:30–12:30 (3 hours) Marks: 50

Instructions:

- This question paper consists of a total of 9 questions worth 50 marks. All questions must be answered; there are no optional ones.
- Make sure to write you name and roll number on the main answer-sheet *and* on all the additional sheets that you use.
- If you need to use the restroom, ask for a toilet pass from the invigilators.
- Attendance will be taken during the exam. Please sign next to your name on the attendance sheet.
- **NOTE:** Your answers must be in mathematical notation, instead of english, as much as possible. You must properly define any mathematical terms you write. Your answers must be precise, complete and to the point. Otherwise, marks will be deducted.
- **NOTE:** Your handwriting must be clear and legible. Answers written in poor and hard-to-read handwriting will be marked as incorrect.
- **WARNING:** CMI's academic policy regarding cheating applies to this exam. Any form of copying or use of electronic devices is strictly prohibited. Violation will result in an automatic F grade, immediate exclusion from placement activities, and potential expulsion without exception.

1. Let X be the space of all input instances and D is a distribution on X. Let $C \subseteq X$ be a concept class, and let $S \subset X$ be a sample set drawn by sampling from D. Let h be a hypothesis. Then describe what is the training-error and true-error of h? *(5 marks)*

2. Explain why we need gated Recurrent Neural Networks such as LSTM. Why do they perform better than standard RNNs? *(5 marks)*

3. Suppose that you are designing a neural network for some task, such as image classification. You have 3 choices for the depth of the neural network:

   (a) Shallow, i.e. around 5 layers

   (b) (Moderately) Deep, i.e. around 50 layers

   (c) Very Deep, i.e. around 500 layers

   Which of these networks would be the best choice, considering both training and inference? Please justify your answer *(5 marks)*

$v_{\pi^*}(s) = \sum_{s', r} p(s', r | s, a) (r + \gamma v_{\pi^*}(s'))$

$q_{\pi^*}(s, a) = \sum_{s', r} p(s', r | s, a) (r + \operatorname{argmax}_a \gamma q_{\pi^*}(s', a))$

4. Recall that we studied 2D convolutions in class. Can you describe a 1D convolution of kernel size 5 with 4-channel input. Write down the complete mathematical equation for this convolution. *(5 marks)*

5. Describe the training of Recurrent Neural Networks using Back-Propagation Through Time. *(5 marks)*

6. Describe the Bellman Optimality Equations for an Markov Decision Process *(5 marks)*

7. Explain the effects of the learning rate during the training process when it is *(5 marks)*

   (a) too high

   (b) too low

8. Describe the loss-function of the Discriminator and the Generator of a GAN. *(5 marks)*

9. Suppose that you have to build a neural network that takes black and white images as input, and produces color images as output. This can be used to turn old-black and white images to color images. Describe:

   (a) How will you construct the training set for this task? *(3 marks)*

   (b) The architecture of the neural network you will use for this task. *(3 marks)*

   (c) The loss function *(2 marks)*

   (d) You observe that, when your neural network is applied to old black-and-white photos not in the training set, sometimes the output image has errors which could of many different types. Describe one concrete reason for this. *(2 marks)*

Please give specific and detailed answers.
