AML-2022 MidSemester Exam

Chennai Mathematical Institute

27th September 2022, 2PM to 4PM — 50 marks

The following questions carry 5 marks each.

1. What are the training error, test error and generalization gap?
2. What are activation functions in neural networks? Is $f(x) = 93x + 7$ a suitable activation function?
3. What is a suitable loss function for classification problems, *Cross Entropy Loss* or Mean Squared Error? Please justify your answer.
4. Explain how recurrent neural networks are trained using backpropogation.
5. Can CNNs handle small rotations and translations in the input image? Please justify your answer.
6. What is the output of the Encoder of a Variational Autoencoder when the latent space has dimension 20? What is the dimension of the input to the decoder?

The following questions carry 10 marks each.

7. Consider a neural network that takes 30×30 3-channel color images as input and in the first layer there are 10 convolution filters each of kernel size 3×3 and stride 1.
   (a) What is the dimension of the output of the first layer?
   (b) How many parameters (weights and biases of the neurons) are there in the first layer?
   (c) What would be the output dimension of the first layer if the stride was 2?
   (d) What would be the output dimension of the first layer if the kernel size was 2×2?
8. Consider the step-function $f_\alpha(x) = \alpha$ if $x \geq \alpha$ else $f_\alpha(x) = \alpha - 1$. Suppose that we use f_1 as our activation function in a neural network for all neurons. Then what is the effect of this choice of activation function on the training of this network? Specifically what happens to the gradients during backpropogation. Please explain in detail.

AML-2022 Final Exam

Chennai Mathematical Institute

23th November 2022, 2PM to 5PM

100 marks

The following questions carry 5 marks each.

1. What is overfitting? What happens to training error and test error in the case of overfitting?
2. What is batch normalization?
3. What happens to the loss function value of the generator and discriminator of a GAN as we train the network?
4. What is Leaky-ReLU? How is it different from ReLU?
5. Why do we use regularization terms in the loss function?
6. What is the "de-convolution" operation? Where is it applied? How is it related to the Transposed Convolution operation?
7. What are Markov Decision Processes?
8. What are the state-value function and state-action-value functions in Reinforcement Learning? Can you always obtain one of these functions from the other one?
9. What is the difference between on-policy and off-policy reinforcement learning?
10. What is VC-dimension? How is it related to the PAC Learning theorem?
11. What are the pros and cons of using a large learning rate in training neural networks? What are the pros and cons of using a small learning rate?
12. Consider the setting of Bayesian Optimization, where $\tilde{f}$ be a statistical model of a function $f : \mathbb{R}^d \to \mathbb{R}$ defined using a gaussian process. Then for a point $x \in \mathbb{R}^d$, what are $f(x)$ and $\tilde{f}(x)$.

The following questions carry 10 marks each.

13. Explain the role of various gates in an LSTM.
14. Explain how multichannel convolution filters works in CNNs. Do we need them when working with black-and-white images? And what about color images?
15. What is the loss function used in Deep Reinforcement Learning? Explain what happens to the loss function value as we train the network. Please justify your answer.
16. Please desribe the loss function used for Variational Autoencoders? Please explain the effect of each term of the loss function on the latent space and output of the autoencoder.