

1. Which of the following regularization in NN (implemented in PyTorch) lead to sparse solutions?
  - 1.1 L1 regularization
  - 1.2 L2 regularization
  - 1.3 Dropout
  - 1.4 Data Augmentation
  - 1.5 None of the above

AC

2. A sparse set of weights in a Deep MLP is preferred:
  - 2.1 it could lead to better generalization
  - 2.2 it is compact and fit in lesser memory
  - 2.3 it has many zeros and lesser amount of operations in forward pass
  - 2.4 it is easy to train when the number of weights/parameters are less
  - 2.5 All the above

ABCDE

3. While re-using a trained network for a new task:
  - 3.1 We always prefer to take the later (towards the end) layer
  - 3.2 We always prefer to take an early(in the beginning) layer

- 3.3 Which layer is more appropriate depends on the tasks.
- 3.4 All the layers are equally useful.
- 3.5 None of the above.

B

- 4. It is believed that adding noise is some sort of regularization.
  - 4.1 Adding noise to the input is useful.
  - 4.2 Adding noise to the output/labels is useful. (for simplicity, assume the task is regression!).
  - 4.3 Adding noise to the weights is useful.
  - 4.4 Higher the noise the better the regularization.
  - 4.5 Lower the noise the better the regularization

ACD

- 5. Consider a problem where we do data augmentation and early stopping.
  - 5.1 With data augmentation, training accuracy is expected to increase.
  - 5.2 With data augmentation, training accuracy may decrease.
  - 5.3 With data augmentation, performance on the validation set is expected to increase.

- 5.4 With data augmentation, the iteration where we do early stop, will increase.
- 5.5 With data augmentation, the iteration where we do early stop, will decrease.

BCE