University of Science and Technology
Communications & Information Engineering Program
CIE 555: Neural Networks and Deep Learning
Spring 2020

Assessment (15 points)

Please upload your answers to Google Classroom before May 1st, 2020

1) (2 points) Which regularization method leads to weight sparsity? Explain
2) (1 point) We discussed how Convolutional Neural Networks leverages parameter sharing to help improve the training of a deep learning model. Explain parameter sharing in CCNs and how they are different from traditional neural networks. Use figures and examples to illustrate your answer.
3) (1 point) Discuss the importance of Batch Normalization in optimizing deep neural networks?
4) (3 points) Consider a CNN that has the following layers:
    - Input layer: one channel of size 128x128x3
    - Convolutional layer: 32 filters with size 9x9 each. Padding is 0 and stride is 1
    - Pooling layer: a 2x2 max-pooling layer with stride 2 and padding 0
    - Convolutional layer: 32 filters with size 5x5 each. Padding is 0 and stride is 1
    - Pooling layer: a 2x2 max-pooling layer with stride 2 and padding 0
    - Fully connected layer: with 5 output neurons

Calculate the total number of weights, number of biases and the dimensions of the feature maps along the network.

5) (3 points) Residual learning is a framework that improves the training process in very deep learning. Discuss residual learning, its main advantage, and some of the architectures that successfully implement it to train deep networks. (you need to mention your references).
6) (1 point) Discuss the advantages of RMSProp algorithm when used in training deep networks.
7) (2 points) Explain the difference between the following design patterns in recurrent neural networks highlighting their applications:
    a. Bi-directional RNNs
    b. Encoder-Decoder Sequence-to-Sequence RNN

Assessment (15 points)

For each design pattern, draw a diagram that illustrates the interaction between the input, hidden units and output

8) (2 points) Why do we need regularization in overcomplete autoencoders? Discuss the regularization techniques in the following autoencoder networks:
    a. Sparse Autoencoders
    b. Denoising Autoencoders
    c. Contractive Autoencoders