

Submission for Deep Learning Exercise 2

Team: dl2022-ryd

Students: Yumna Ali, Rean Fernandes

December 6, 2022

1 Question 2

1.1 In what sense are convolutional neural networks and recurrent neural networks similar? In what sense are they different?

The similarity between CNNs and RNNs is the neural architecture that they use. They both involve at the fundamental level, matrix multiplications with learnable weights and biases (parameter) and use backpropagation to update these parameters.

The difference is that while the CNN can only learn contextual information in the spatial dimensions, RNNs can learn contextual features in data across the time dimension as well. Thus they are more suited to dealing with sequential data like videos and times series forecasting.

1.2 How can one counteract vanishing or exploding gradients in RNNs, allowing to learn long-term dependencies?

Gradient clipping can be used to prevent vanishing/ exploding gradients by letting the values for all gradients lie between a bound. if the values should exceed or fall below this bound, they will automatically be set to the maximum/minimum value of the bound. Another would be to use a more complex architecture like LSTMs, GRUs and others, which while increasing complexity also ensure the long term propagation of memory. Another potential method could be to use a ReLU function instead of a sigmoid activation, but this would only deal with the exploding gradient problem and not the vanishing gradient, which is why we are not so sure of it, but included it nonetheless.