Stevens Institute of Technology

Department of Computer Science
(CS583 - Deep Learning)
Assignment 2

Professor: Jia Xu

Student Name: Dhruv Vaghela
CWID: 20015603
Date: 23rd Dec 2022

Q4. If you input a very long sentence, do you have problem of translation this sentence? If so, what could be potentially the problem? Write down the answer.
Ans. When the input is a very long sentence the gradient converges to a small value as the back-propagation algorithm moves through the network. Due to this, the previous layers learn at a slower rate than the upcoming layers. This hampers the model's potential to take long sentences as inputs. As the information gap between the layers increases the gradient shrinks down which is also known as the vanishing gradient problem.

Q5. What model or activation function can you use the improve long-sentence translation? Implement such a model.
Ans. To address the aforementioned problem in Q4, we can use attention-based RNN models which can focus on different parts of the input for each previous output pair since it can be well suited for translating long sentences when compared to no-attention RNN models.
We can also use activation functions like ReLU which can also help in resolving vanishing gradients.