1. **What happens to the gradient if you backpropagate through a long sequence?**

Sol: Gradient contributions from “far away” steps become zero, and the state at those steps does not contribute to what you are learning.

In a word, the gradient will be vanished.

1. **What happens if you increase the number of hidden layers in the RNN model?**

Sol: The depth of the models increased, but the prediction accuracy may increase.