Multiple Layers

Stack multiple LSTM cell layers for added performance.

Chapter Goals:

- Learn how to stack multiple cell layers in an RNN

A. Stacking layers

Similar to how we can stack hidden layers in an MLP, we can also stack cell layers in an RNN. Adding cell layers allows the model to pick up on more complex features from the input sequence and therefore improve performance when trained on a large enough dataset.

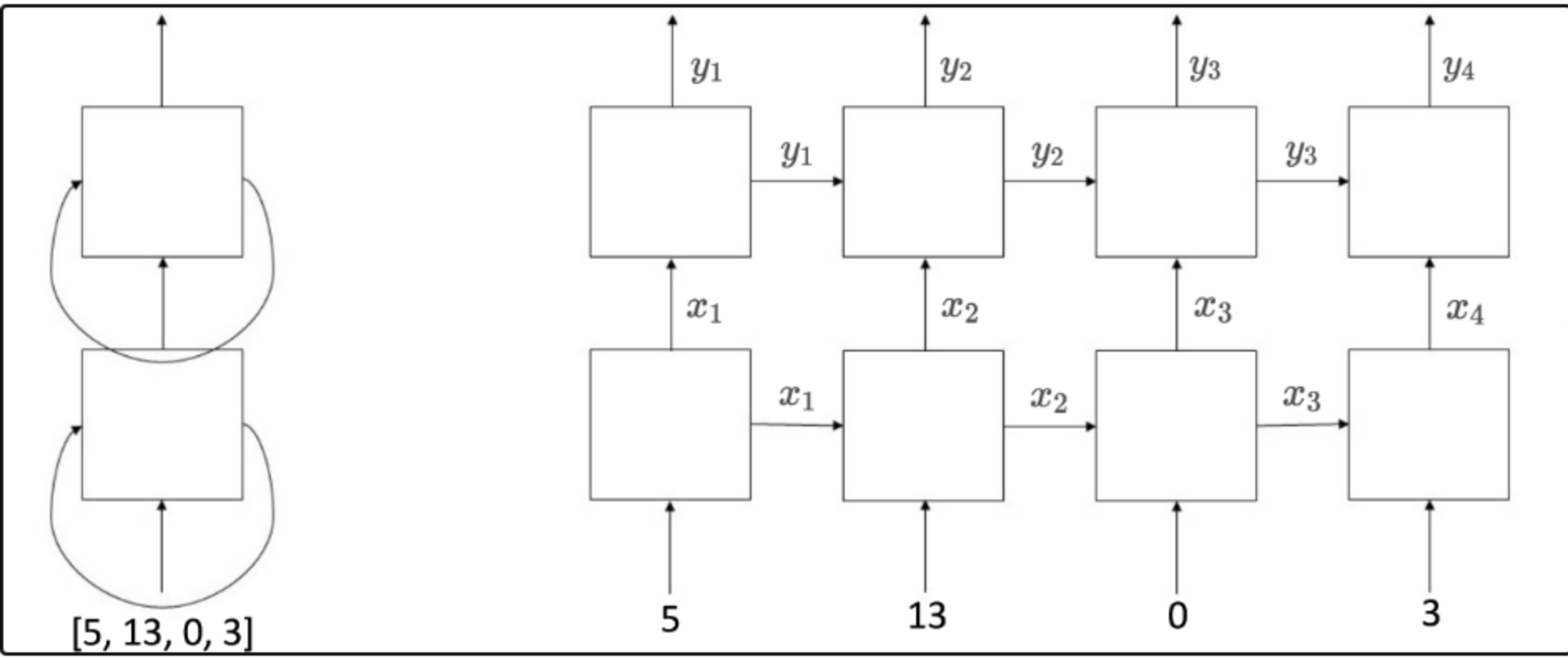

Diagram of an RNN with 2 cell layers. On the left is the rolled RNN, on the right is the unrolled RNN with 4 time steps.

In the diagram above, the RNN contains 2 cell layers (which is just two cells). At each time step, the first cell's output becomes the input for the second cell, and the second cell's output is the overall RNN's output for that particular time step. One thing to note is that the cell states remain separate, i.e. there aren't any recurrent connections between the two cells.

The input in this example is `[5, 13, 0, 3]`, which represents the tokenized sequence for some four word sentence. x_1, x_2, x_3, and x_4 represent the outputs for the first layer. y_1, y_2, y_3, and y_4 represent the outputs for the second layer.

As with all other neural networks, adding layers can improve performance on larger datasets but also run the risk of overfitting the data. This makes regularization techniques such as dropout more important as we increase the size of our model.