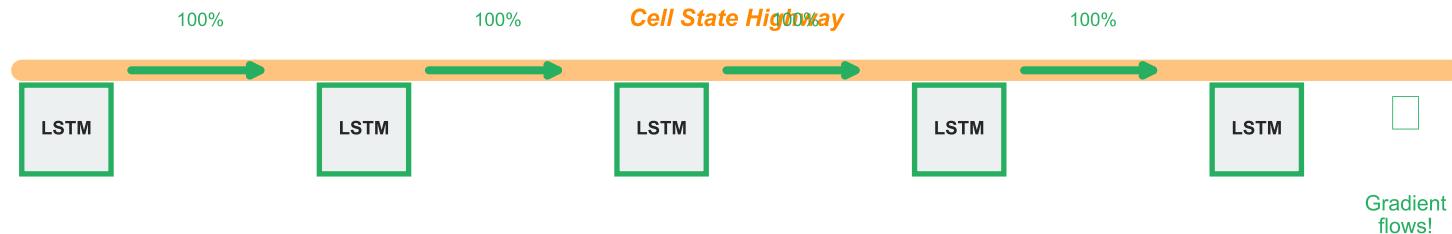


# The Vanishing Gradient Problem

## Standard RNN:



## LSTM:



Key: LSTM uses addition (cell state) instead of multiplication (RNN hidden state)