

什么是长短时记忆网络(LSTM)

RNN有一定的记忆能力，但不幸的是它只能保留短期记忆，在各类任务上表现并不好，那该怎么办？

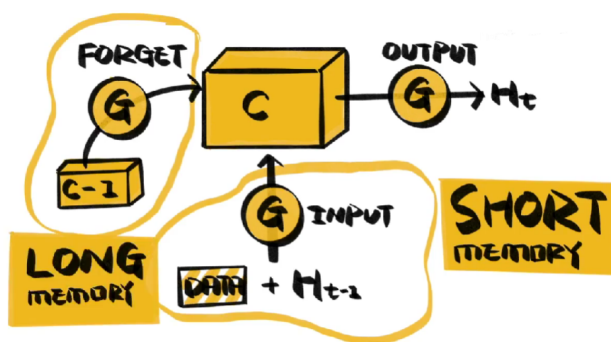
人们将目光投向了自己，人类的记忆是有取舍的，我们不会记住每时每刻发生的所有事，会有选择的留下重要的，丢掉不重要的。

参考这种记忆机制，人们改造了RNN中使用的“小盒子”，并找到了“门”这种机制，门是用来决定信息如何保留的小开关，它的数值在0-1之间，1是完全保留，0是完全舍弃。

新的小盒子上有三个门

1. 其中遗忘门决定了小盒子里要保留多少原有信息，也就是丢掉哪些不重要记忆
2. 输入门决定了当前网络信息有多少要被保存到小盒子里，也就是记住哪些新东西
3. 输出门决定了多大程度的输出小盒子中的信息

经过改造的小盒子，既能通过输入门对当前网络状态有所了解，又能利用遗忘门留下过往的重要信息，这就是**LSTM**-长短时记忆模型。



通过改变小盒子的结构，LSTM还有很多变体，如MGU、SRU。

不过最受欢迎的还是门控循环单元**GRU**。GRU只有两个门，更新门是遗忘门和输入门的结合体，决定丢弃哪些旧信息，添加哪些新信息，重置门决定写入多少上一时刻网络的状态，用来捕捉短期记忆，结构更简洁，计算更高效，效果又与LSTM不相上下，GRU正变得越来越流行。

