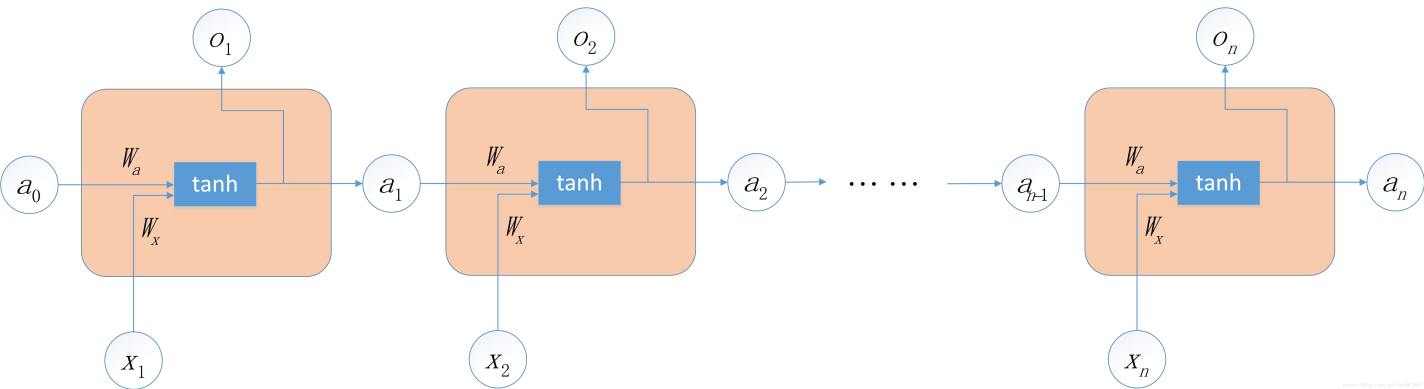


BiLSTM (双向长短时记忆网络)

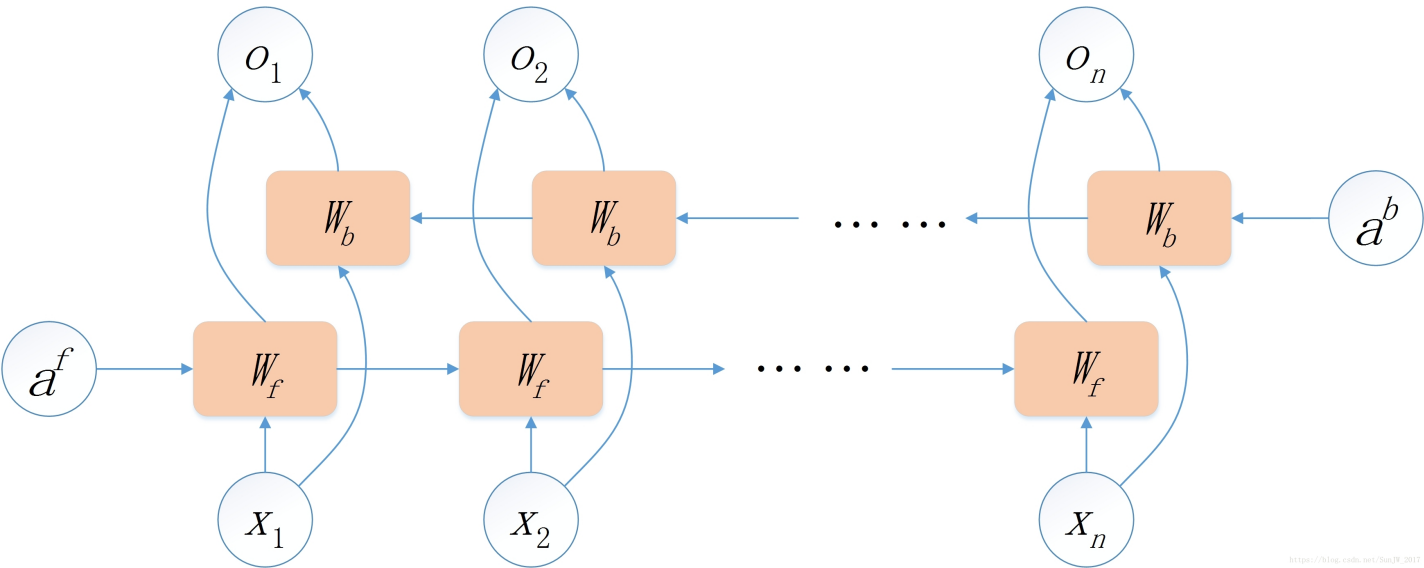
RNN

普通的神经网络不容易处理训练样本输入为连续序列的情况，比如一段语音、一段文字，于是 RNN 就诞生了。RNN 的运行机制如下图：



从图中看出，与普通神经网络不同的一点在于，神经元接受两个值：一个是当前时刻的输入 x_t ，一个是上一个神经元的输出 a_{t-1} ，通过将前一时刻的运算结果添加到当前的运算中，从而实现“考虑上文信息”的功能。

如果不仅要考虑上文的信息，还要考虑下文的信息，可以使用 BiRNN，BiRNN 正反向分别计算，然后将正反向结果堆叠，生成最终的结果：



RNN 存在的问题

对某些简单的问题，可能只需要最后输入的少量时序信息即可解决。但对于某些复杂问题，可能需要更早的

一些信息，而 RNN 难以记忆间隔太久的输入信息，这时候就需要用到 LSTM 了。

LSTM

LSTM 最核心的是引入了“门”的概念，门实际上是一种全连接层，它的输入是一个向量，输出是一个0到1之间的实数向量。假设 W 是门的权重向量， b 是偏置项，那么门可以表示为：

$$g(x) = \sigma(Wx + b)$$