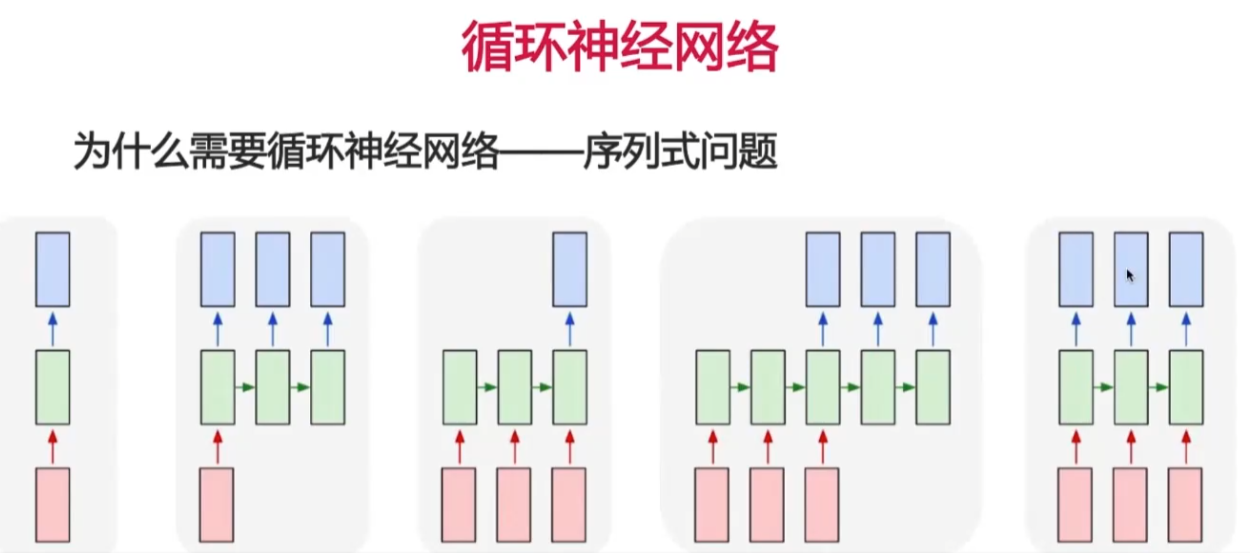
# 了解 LSTM 网络

<http://colah.github.io/posts/2015-08-Understanding-LSTMs/>

cnn有很多变体和结构，但是rnn只有几种。

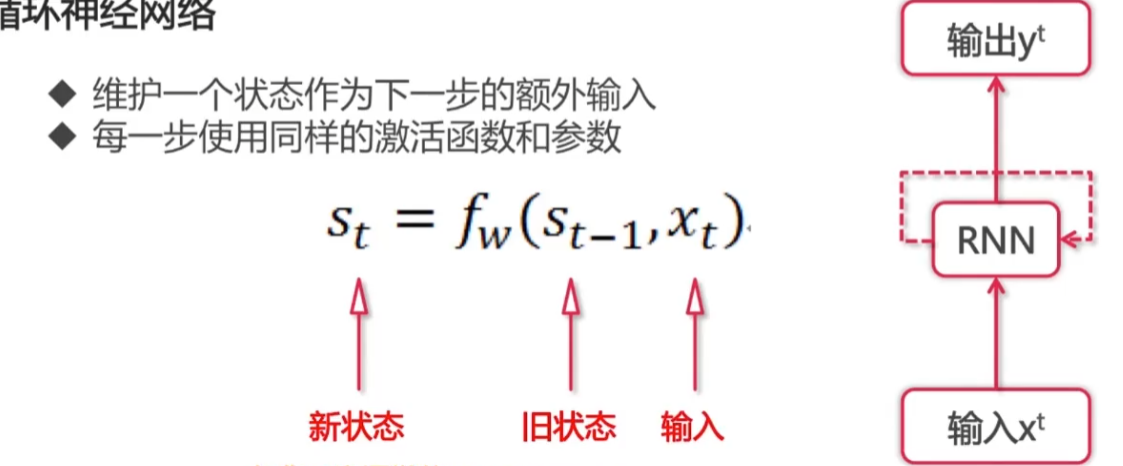
Cnn只适合固定长度的数据，不固定长度的数据不能够实现。

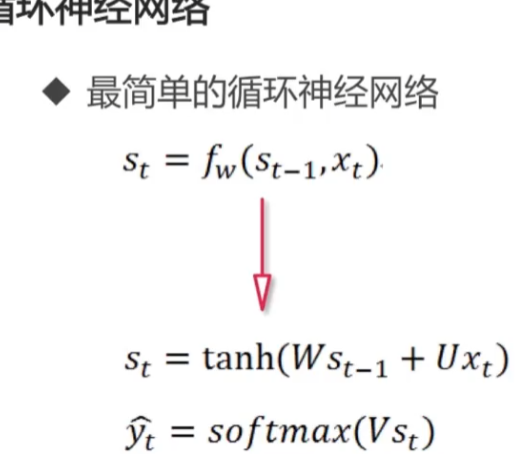


Rnn可以处理cnn处理不了的序列数据。

维护一个状态作为下一步的额外输出

每一步使用同样的激活函数和参数

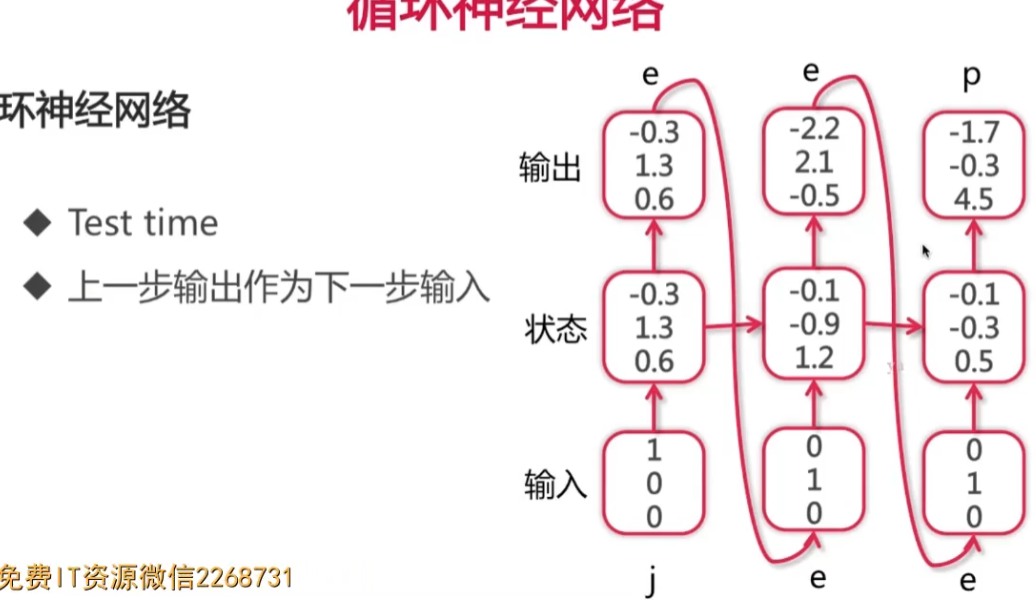




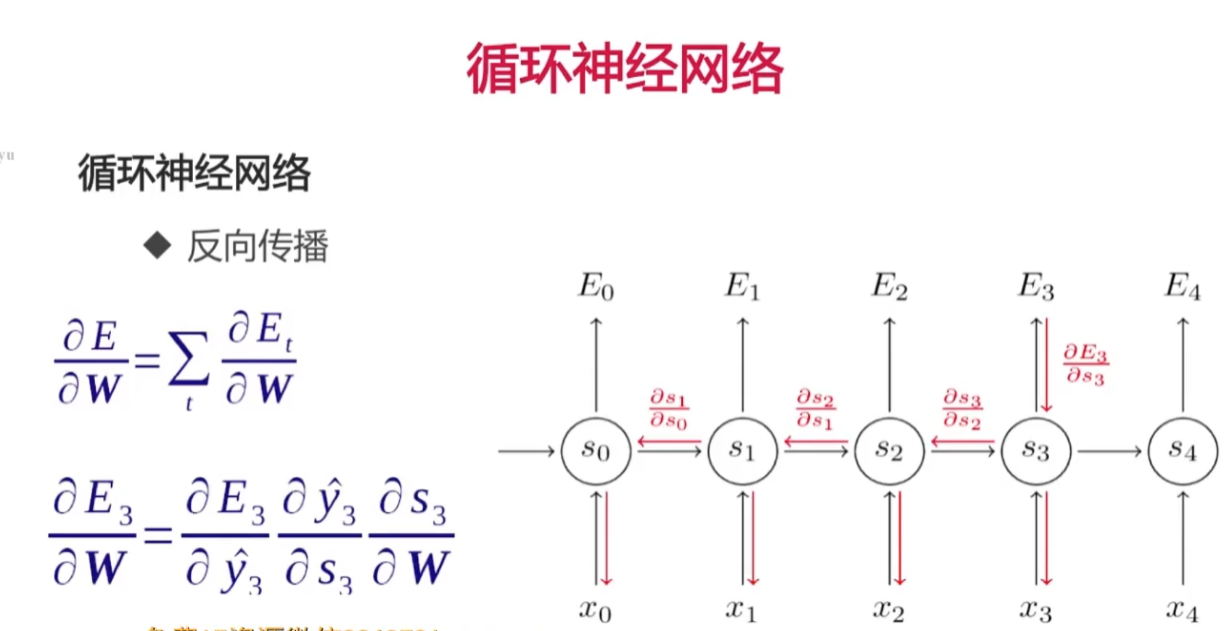
Softmax是一个多分类的计算公式

什么是字符语言模型，就是我给定一个上下文，他就能预测下一个字符和词是什么。

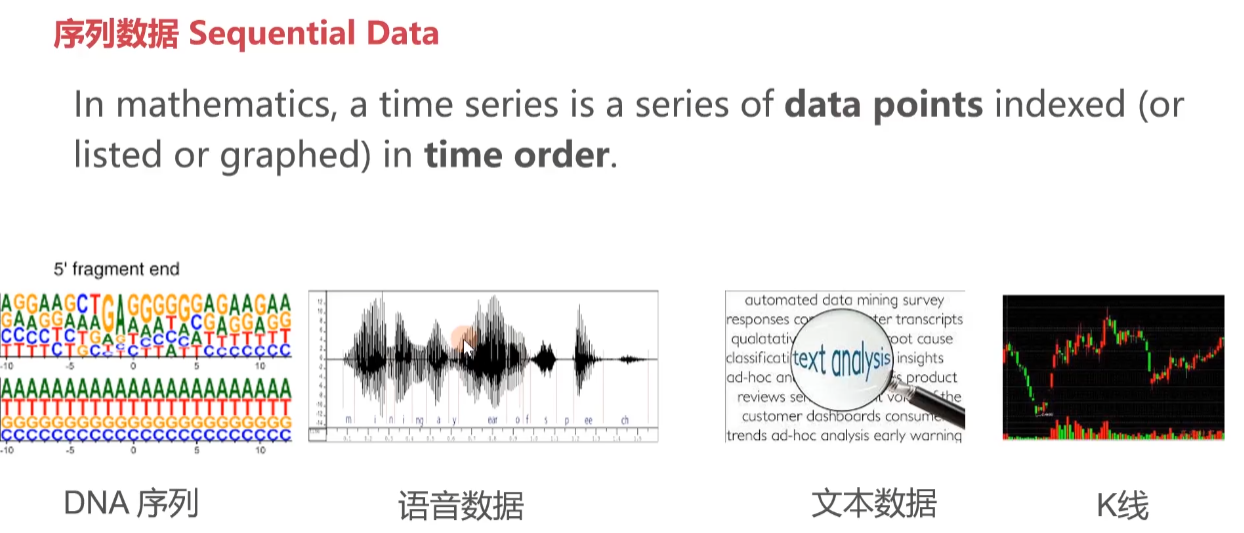
最后一个输入e对p的影响最大，j对p的影响是最小的



反向传播如何计算梯度

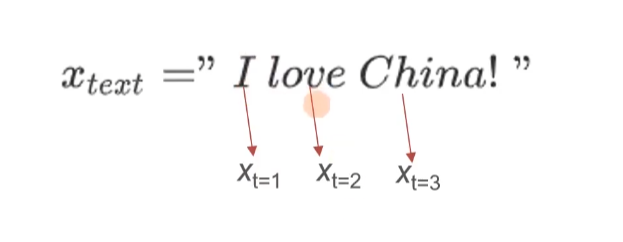


什么是序列数据?

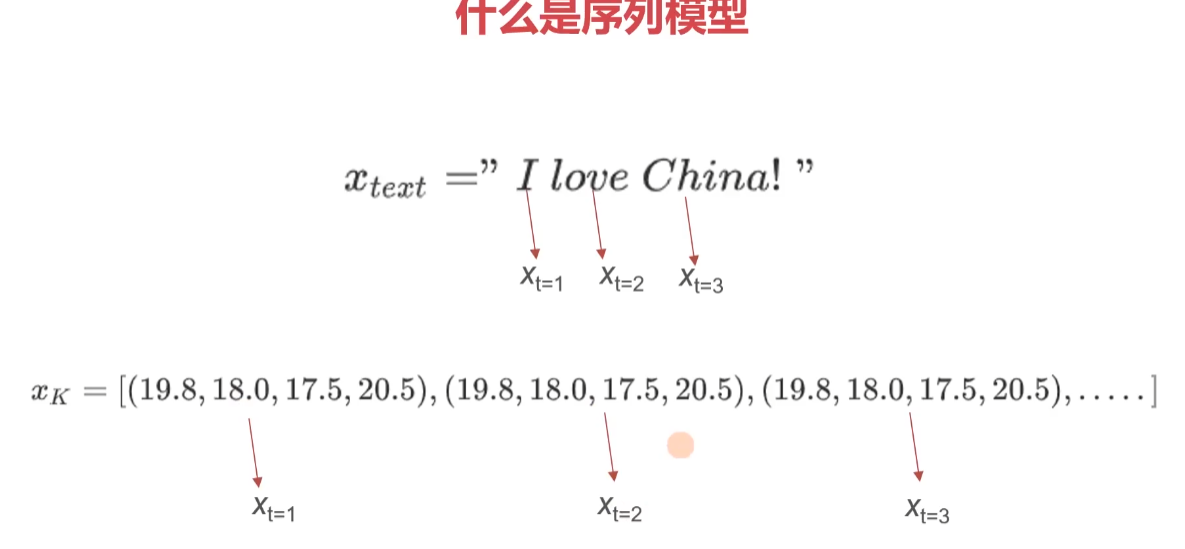


什么是时间步？

可以是一个个字母，也可以是一个个字符

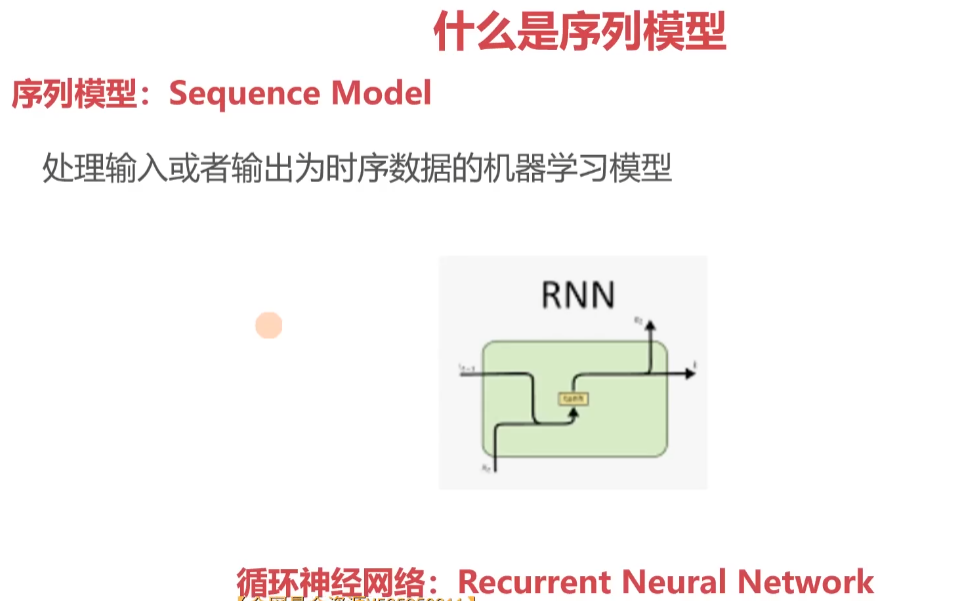


下面相当于一个词有四个通道的意思，相当于颜色有四个通道的意思



使用cnn还是rnn，取决你关心数据的结构特征，还是时序特征





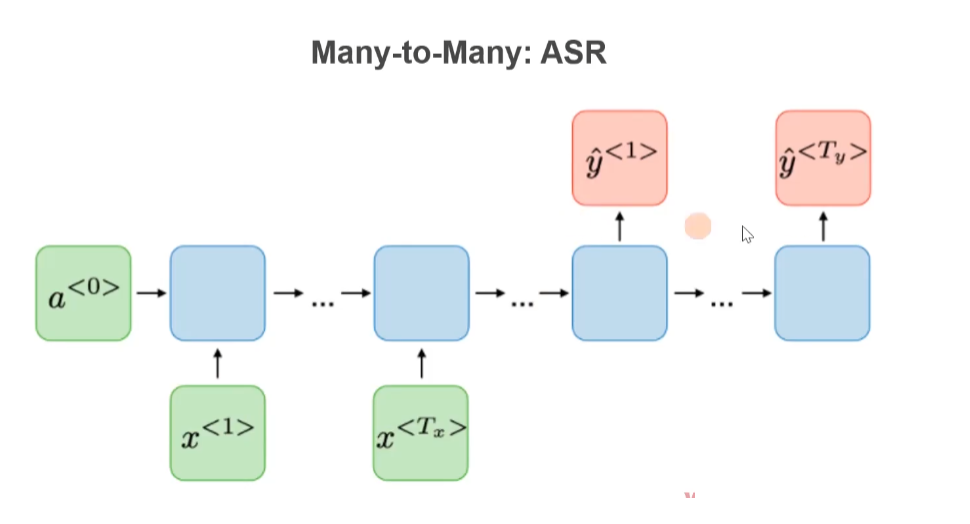
CNN就是固定长度，固定输出。

RNN可以处理什么问题？

可以解决1对1的问题。

可以解决1对多的问题，例如输入摇滚乐，输出一系列的音乐。

下面是rnn解决多对多的问题，例如语音识别



循环神经网络的公式

下图ht-1承载了前面所有时间步的时序特征，加上当前步的数据特征

