



循环神经网络

作者: Calvin

QQ: 179209347

Mail: 179209347@qq.com

介绍

笔记简介:

- 面向对象: 深度学习初学者
- 依赖课程: **线性代数, 统计概率**, 优化理论, 图论, 离散数学, 微积分, 信息论

知乎专栏:

<https://zhuanlan.zhihu.com/p/693738275>

Github & Gitee 地址:

https://github.com/mymagicpower/AIAS/tree/main/deep_learning

https://gitee.com/mymagicpower/AIAS/tree/main/deep_learning

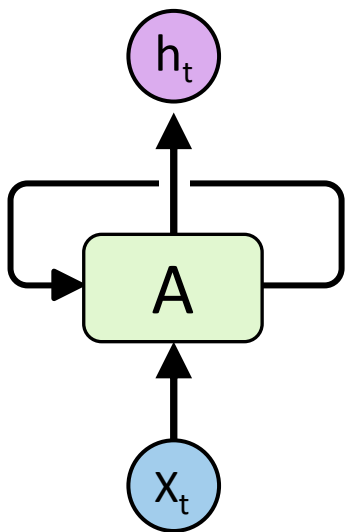
* 版权声明:

- 仅限用于个人学习
- 禁止用于任何商业用途

简单循环神经网络 (Simple Recurrent Network , SRN)

简单循环神经网络由一个循环单元组成，通常使用 tanh 或者 ReLU 等激活函数。

在每个时间步，循环神经网络接收输入 \mathbf{x}_t 和前一个时间步的隐藏状态 \mathbf{h}_{t-1} ，然后计算当前时间步的隐藏状态 \mathbf{h}_t 。这个隐藏状态包含了网络对过去信息的记忆，可以捕捉到序列中的长期依赖关系。



简单循环神经网络的计算过程可以用以下公式表示：

$$z_t = U\mathbf{h}_{t-1} + W\mathbf{x}_t + b$$

$$\mathbf{h}_t = f(z_t)$$

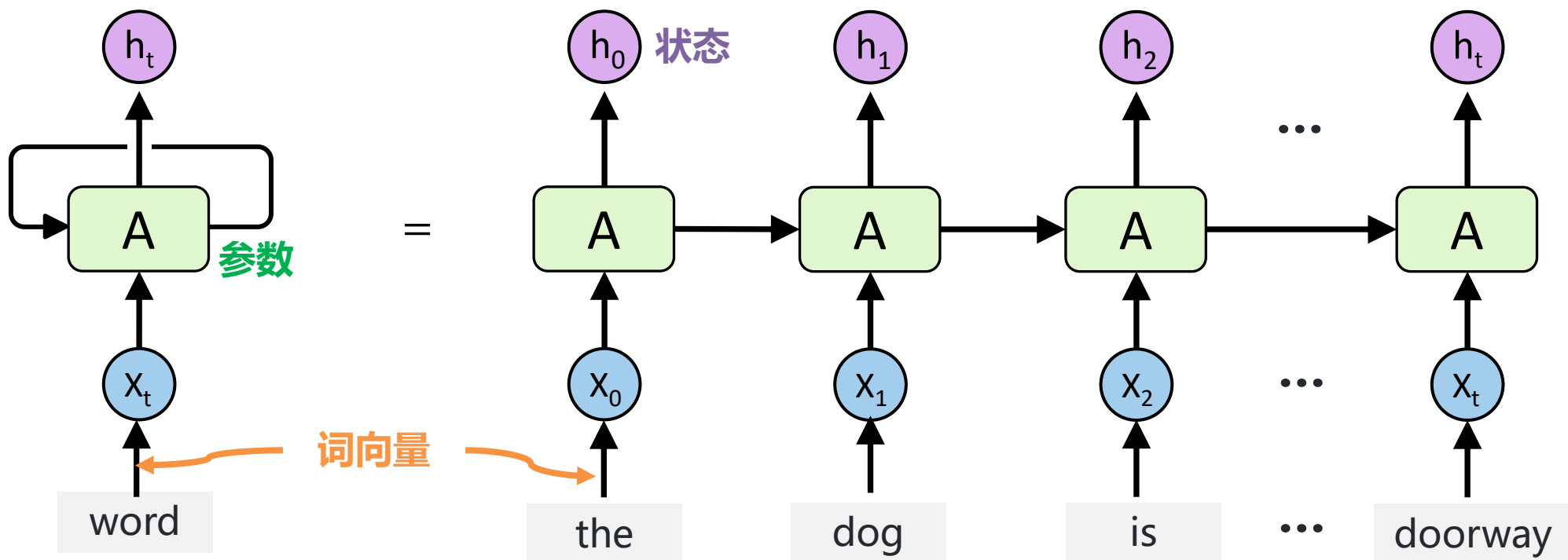
$$= f(U\mathbf{h}_{t-1} + W\mathbf{x}_t + b)$$

- z_t 为隐藏层的净输入
- U 为状态-状态权重矩阵
- W 为状态-输入权重矩阵
- b 为偏置
- $f()$ 是非线性激活函数

简单循环神经网络 – 工作原理

将左侧的图按照时间线展开：

在 "The dog is sitting at the doorway. " 这个例子中，通过下面这张图我们可以看出，循环神经网络前面的结果对后面产生了影响。



简单循环神经网络结构的展开

循环神经网络的通用近似定理

循环神经网络的通用近似定理 [Haykin,2009]:

如果一个完全连接的循环神经网络有足够数量的 sigmoid 型隐藏神经元, 它可以以**任意的准确率去近似任何一个非线性动力系统**。

$$s_t = g(s_{t-1}, x_t)$$

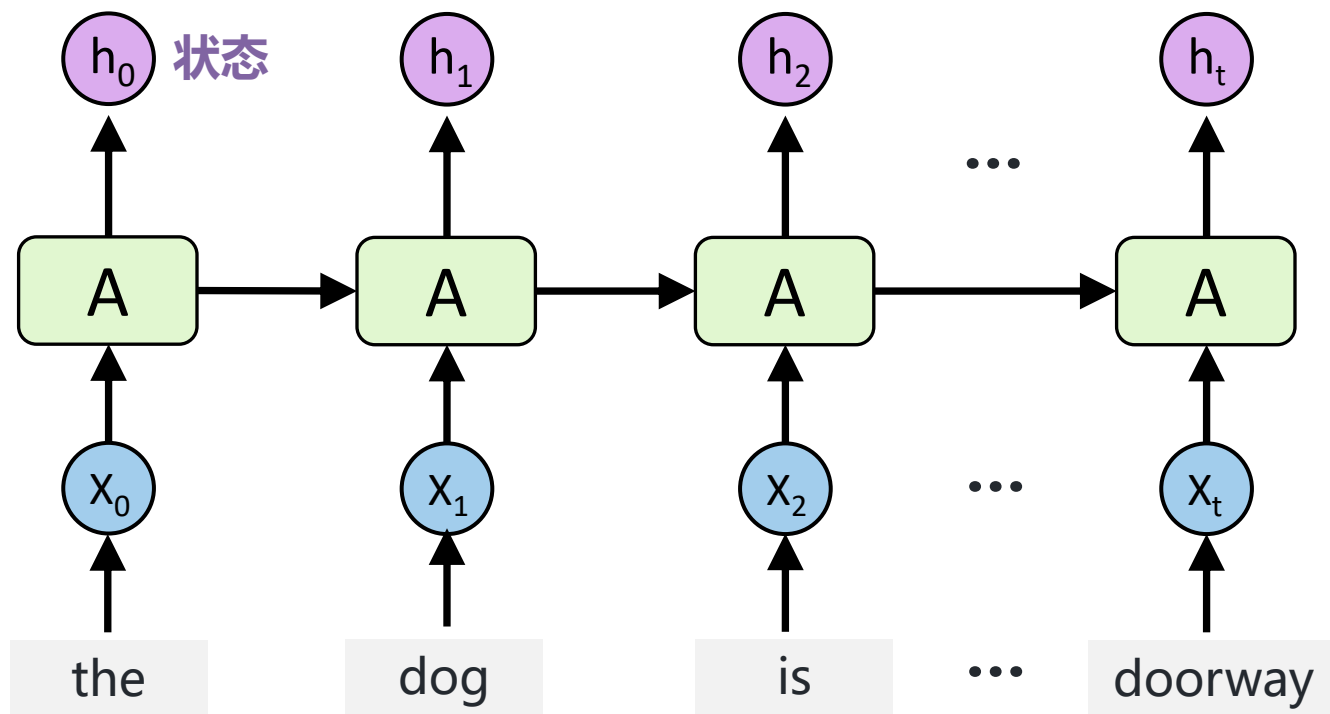
$$y_t = o(s_t)$$

- s_t 为每个时刻的隐状态
- x_t 是外部输入
- $g(s_{t-1}, x_t)$ 是可测的状态转换函数
- $o(s_t)$ 是连续输出函数, 并且对状态空间的紧致性没有限制

简单循环神经网络 – 存在的问题

简单循环神经网络（RNN）一些主要问题包括：

- 梯度消失或梯度爆炸
- 长期依赖性
- 参数共享
- 计算效率
- 难以并行化





Thank

You