



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

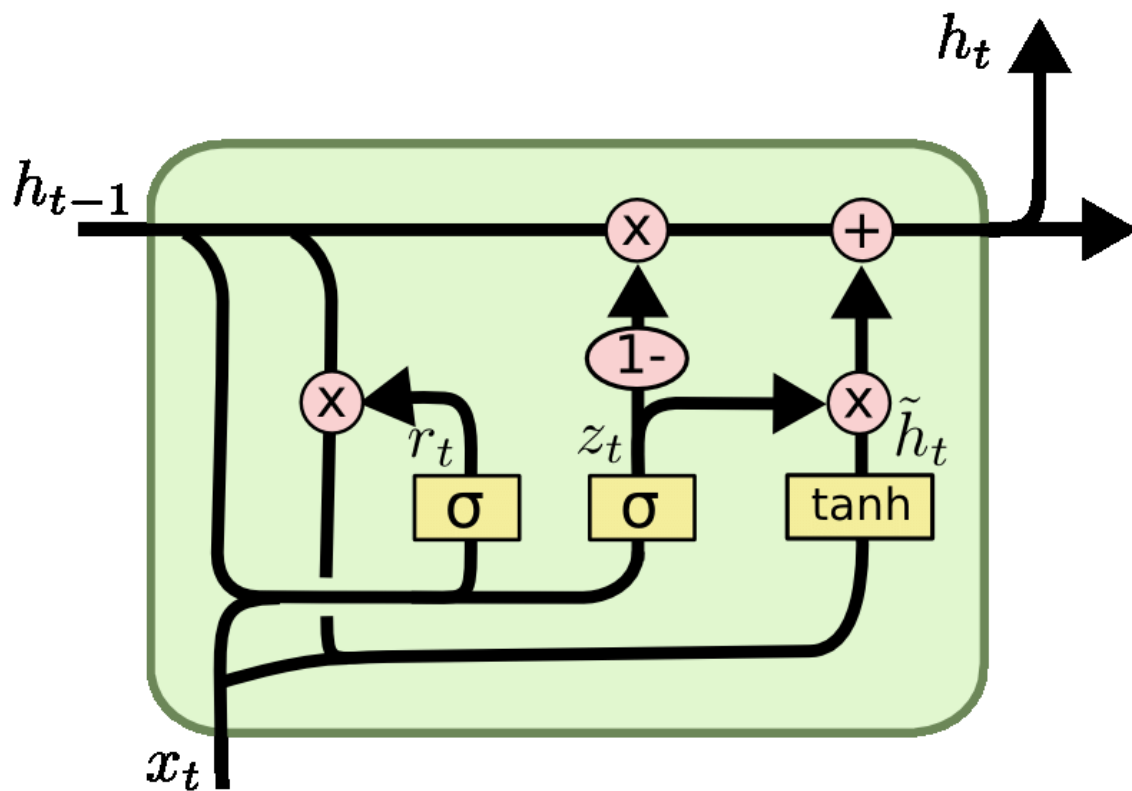
https://gitee.com/mymagicpower/AIAS/tree/main/deep_learning

* 版权声明:

- 仅限用于个人学习
- 禁止用于任何商业用途

门控循环单元神经网络

门控循环单元（Gated Recurrent Unit, GRU）是一种常用的循环神经网络（RNN）变体，旨在解决传统RNN中的梯度消失问题和长期依赖问题。GRU于2014年提出，是LSTM（长短期记忆网络）的一种简化版本。

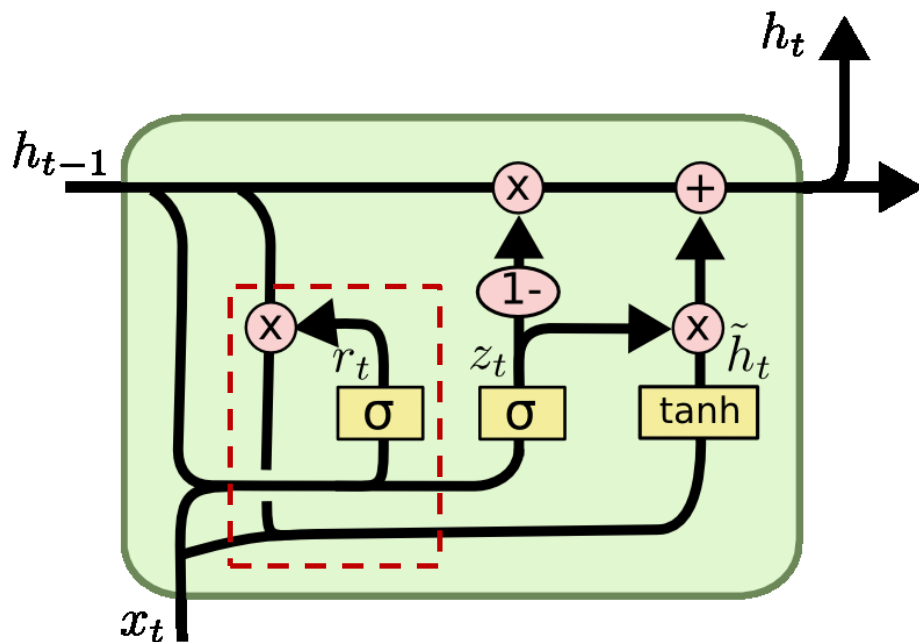


GRU具有以下几个关键部分：

- 重置门（Reset Gate）
- 更新门（Update Gate）
- 当前候选值（Candidate Value）
- 隐藏状态（Hidden State）

门控循环单元神经网络 - 重置门

重置门 (Reset Gate)：帮助网络决定在当前时间步应该忽略多少过去的记忆。重置门的值在0到1之间，0表示完全忽略过去的记忆，1表示完全保留过去的记忆。



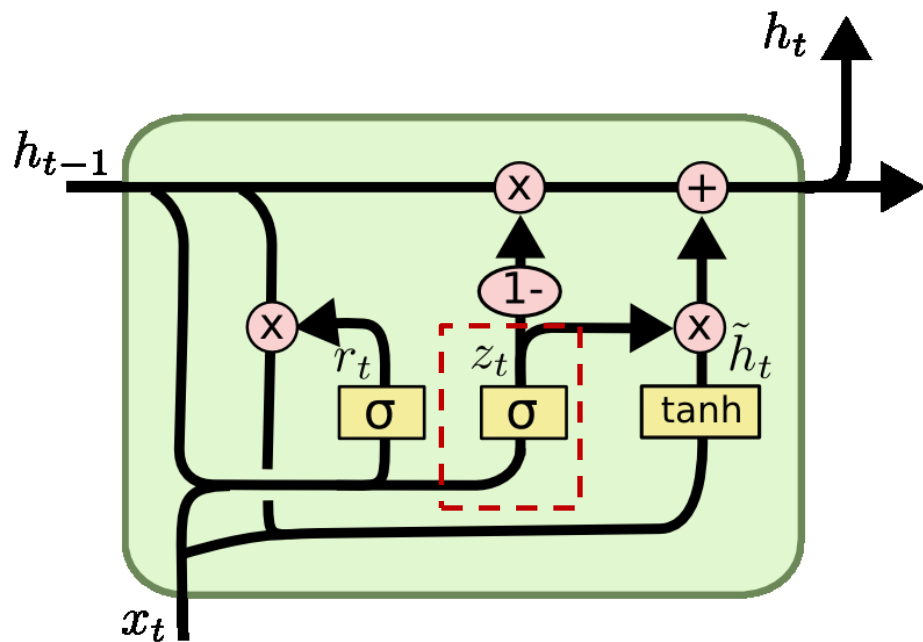
计算公式如下：

$$r_t = \sigma(W_r x_t + U_r h_{t-1} + b_r)$$

- $\sigma()$ 为Sigmoid函数
- h_{t-1}, x_t 为输入数据
- W_r, U_r 为重置门的权重
- b_r 为重置门的偏置

重置门用于控制保留多少之前的数据，使用 r_t 代表重置门，当 $r_t = 0$ 时，表示遗忘之前的所有数据，重设为目前输入数据的状态。

更新门 (Update Gate) : 控制有多少过去的记忆应该被保留。更新门的值在0到1之间, 0表示完全忽略过去的记忆, 1表示完全保留过去的记忆。



更新门用于控制数据的更新与保留，使用 z_t 代表更新门，当 $z_t = 0$ 时，表示保留之前的所有数据。

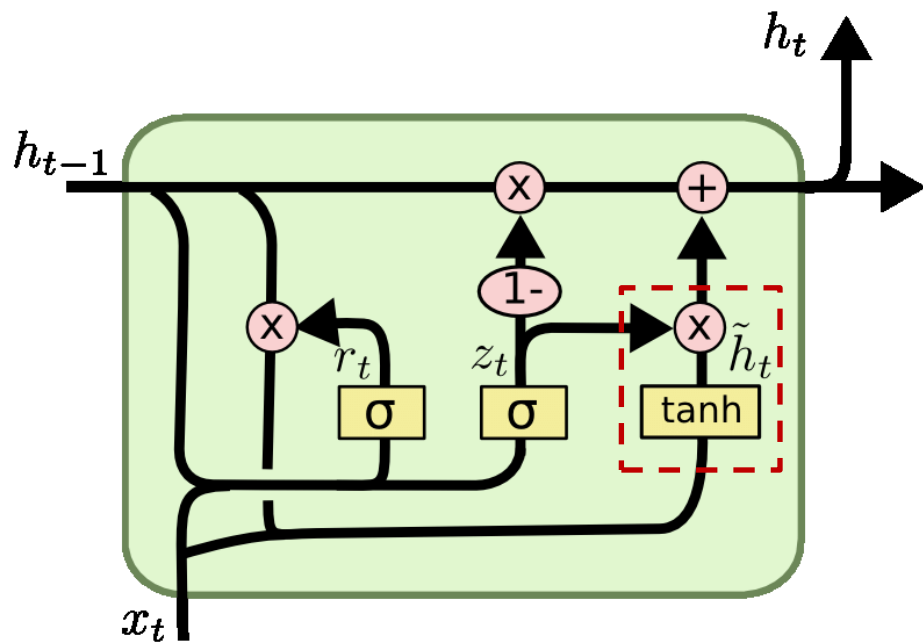
计算公式如下：

$$z_t = \sigma(W_z x_t + U_z h_{t-1} + b_z)$$

- $\sigma()$ 为Sigmoid函数
- h_{t-1}, x_t 为输入数据
- $W_z U_z$ 为重置门的权重
- b_z 为重置门的偏置

门控循环单元神经网络 – 当前候选值

当前候选值 (Candidate Value)：根据当前输入和过去的记忆计算出的候选值，用于更新当前时间步的记忆。



计算公式如下：

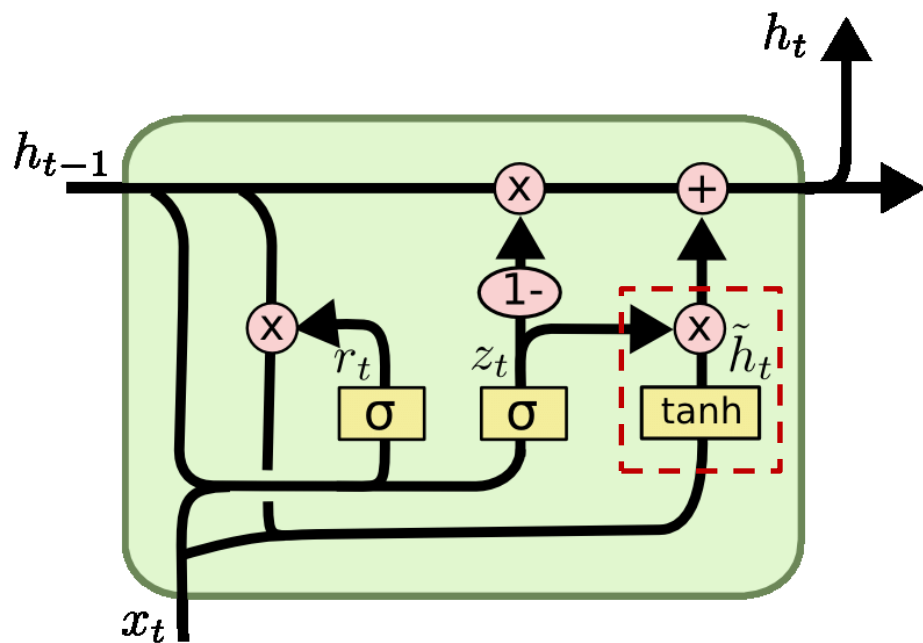
$$\tilde{h}_t = \tanh (W x_t + r_t \odot U h_{t-1} + b)$$

- r_t, h_{t-1}, x_t 为输入数据
- W, U 为可学习的参数
- b 为偏置

候选数据 \tilde{h}_t 其输入数据是将重置门所保留下来的数据 r_t 和目前的输入数据合并成 $[r_t h_{t-1}, x_t]$ 向量。

门控循环单元神经网络 – 隐藏状态

隐藏状态 (Hidden State)：GRU的输出，包含了当前时间步的信息。



最后，计算出GRU单元的输出数据 h_t 。

计算公式如下：

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t$$

- \odot Hadamard积运算符，也就是按元素乘积
- $(1 - z_t) \odot h_{t-1}$ 为之前保留的数据
- $z_t \odot \tilde{h}_t$ 为更新的数据



Thank

You