

长短时记忆网络与门控循环单元

目录

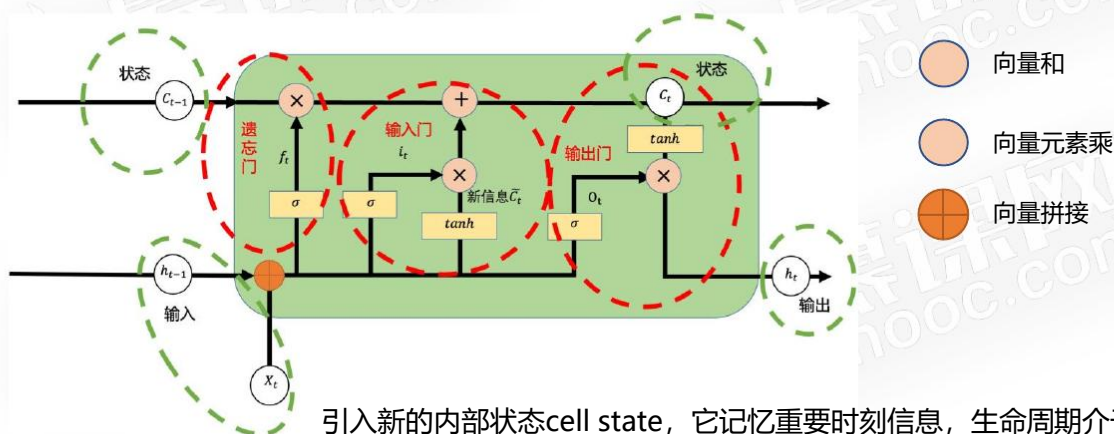
◆ 长短时记忆网络

◆ 门控循环单元

长短时记忆网络

Long Short-Term Memory Network

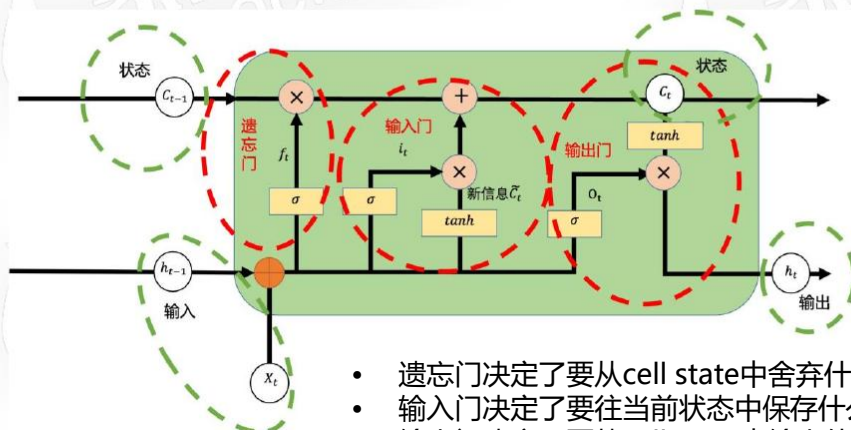
- ◆ 简称LSTM，主要是为了解决RNN训练过程中的长距离依赖问题



引入新的内部状态cell state，它记忆重要时刻信息，生命周期介于长期与短期记忆之间。cell state状态更新是缓慢的，隐藏层h的更新是迅速的。

Long Short-Term Memory Network

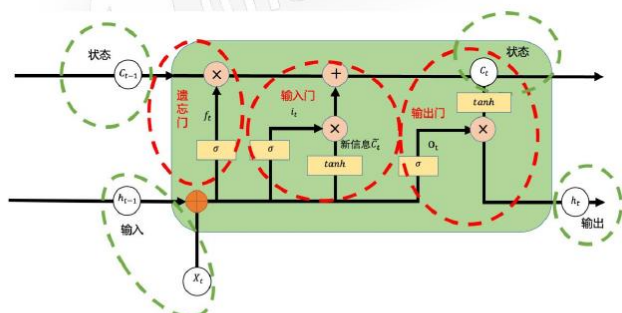
- ◆ 包含了三个门控单元，遗忘门、输入门与输出门，决定每一个时刻要舍弃什么旧的信息以及添加什么新的信息



- 遗忘门决定了要从cell state中舍弃什么信息
- 输入门决定了要往当前状态中保存什么新的信息
- 输出门决定了要从cell state中输出什么信息

Long Short-Term Memory Network

- ◆ 门控单元的计算



遗忘门: ()

上一个状态需要遗忘的信息:

输入门:

新信息: ()

更新信息:

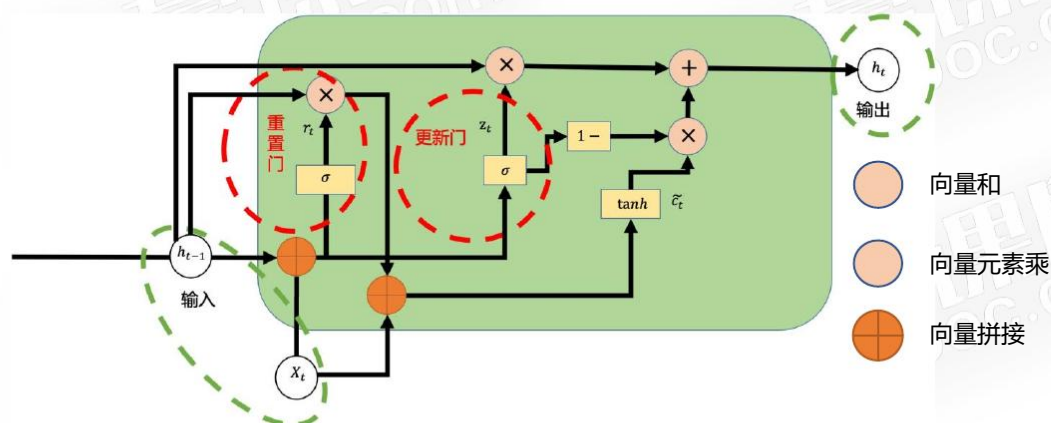
输出门: ()

输出状态: ()

门控循环单元

Gated Recurrent Unit(GRU)

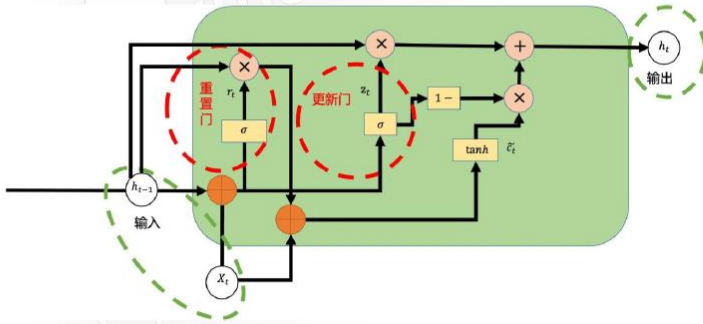
◆ 对LSTM结构的简化，由更新门和重置门组成



- 更新门决定了要从历史信息中保留多少信息，捕捉长程依赖关系。
- 重置门用于控制候选状态是否依赖于上一状态，可以用来丢弃与预测无关的历史信息。

Gated Recurrent Unit(GRU)

◆ 门控单元的计算



更新门: ()

重置门:

新信息: ()

输出:

下次预告：卷积神经网络基础