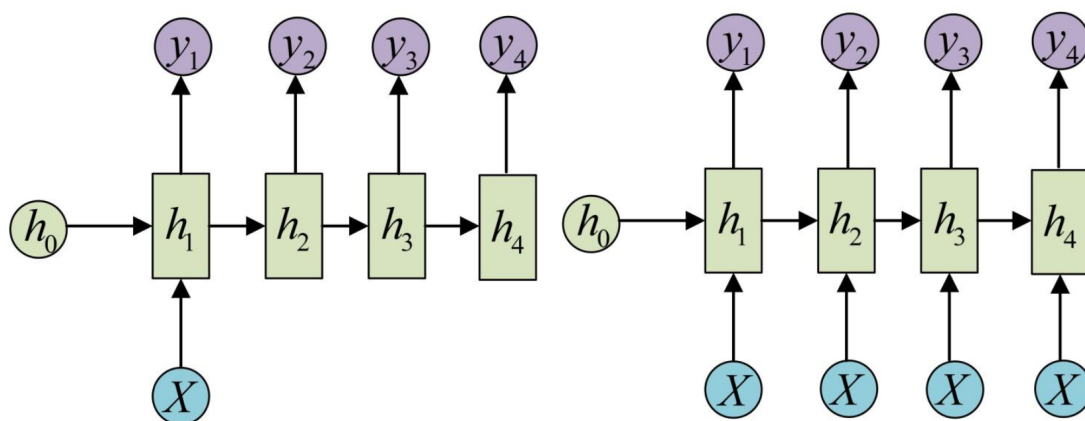


## 解释RNN、LSTM、GRU模型

- RNN (循环神经网络)：RNN 的核心思想是利用序列信息，它通过在网络中引入循环来实现这一点。对于每个时间步的输入，RNN 不仅会考虑当前输入，还会考虑前一个时间步的输出（隐藏状态）。然而，传统的 RNN 在处理长距离依赖时遇到了挑战，比如梯度消失或爆炸问题。

如下图， $y_1 = \text{act}(vh_1 + c)$ ,  $h_1 = f(ux_1 + wh_0 + b)$ ，以此类推



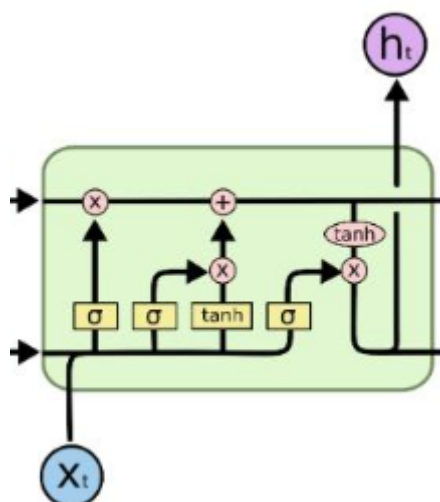
- LSTM (长短期记忆网络)：为了解决传统 RNN 的局限性，LSTM 引入了一种称为“门控”的机制。LSTM 单元包括三个门：遗忘门（决定丢弃哪些信息）、记忆门（更新细胞状态）和输出门（决定下一层接收的信息）。这些门使用 sigmoid 函数和点乘操作来控制信息流，使得 LSTM 能够有效地学习长期依赖关系。

如下图，上方的横线为长期记忆  $c_{t-1} \rightarrow c_t$ ，下方为短期记忆  $h_{t-1} \rightarrow h_t$

遗忘门根据  $h_{t-1}$  和  $x_t$  决定长期记忆的保留率， $f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$ ，用  $f_t$  去乘以  $c_{t-1}$  来实现遗忘

记忆门根据  $h_{t-1}$  和  $x_t$  决定记住当前多少，当前信息  $i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$ ，记忆率  $C'_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c)$ ，两者相乘加到长期记忆上

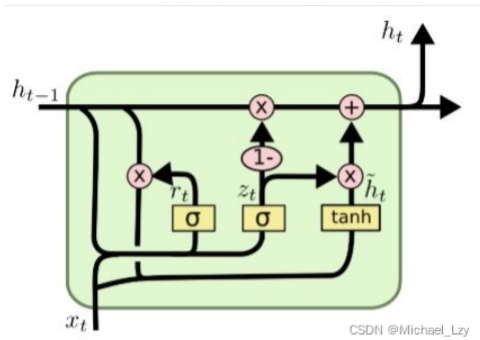
输出门负责给出当前神经元的输出，输出原始信息  $o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$ ，输出率  $\tanh(c_t) = \tanh(c_{t-1} \cdot f_t + i_t \cdot C'_t)$ ，于是输出  $h_t = o_t \cdot \tanh(c_t)$



- GRU (门控循环单元)：GRU 是 LSTM 的一种变体，旨在简化模型结构同时保留 LSTM 的关键特征。GRU 将 LSTM 中的遗忘门和输入门合并成一个单一的“更新门”，并结合了隐藏状态和细胞状态。这不仅减少了参数数量，还加速了训练过程。

下图中

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t])$$
$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t])$$
$$h'_t = \tanh(W_h \cdot [r_t * h_{t-1}, x_t])$$
$$h_t = (1 - z_t) * h_{t-1} + z_t * h'_t$$



## 叙述诗歌生成过程

1. **预处理**：首先需要对文本数据进行预处理，包括分词、建立词汇表以及将词语转换为索引。
2. **构建模型**：根据任务需求选择合适的模型架构LSTM或RNN，并定义相应的层和参数。
3. **训练模型**：使用大量文本数据来训练模型，使其能够学习到语言模式和语法规则。在训练过程中，模型尝试基于先前的词预测下一个词。
4. **生成诗歌**：
  - 从指定的起始词开始，将其转换为模型可以接受的形式——嵌入向量。
  - 使用训练好的模型预测下一个词的概率分布，并从中选取最可能的一个作为新的输入。
  - 重复此过程，直到生成满足条件的诗句长度或者遇到特定的终止符为止。
  - 最终，将生成的索引序列转换回单词序列以形成完整的诗歌

```
epoch 29 batch number 346 loss is: 6.708414077758789
prediction [10, 5, 15, 15, 32, 0, 272, 0, 10, 95, 30, 432, 91, 18, 104, 1, 10, 18, 16, 21, 12, 0, 94, 0, 16, 7, 10, 5, 9, 21, 22, 1, 30, 416, 18, 18, 0, 0,
0, 0, 83, 18, 83, 16, 28, 12, 71, 1, 10, 178, 10, 19, 83, 0, 272, 0, 10, 233, 9, 5, 10, 1, 94, 1, 10, 74, 18, 0, 30, 23, 170, 0, 10, 41, 33, 21, 10, 10, 1
4, 1, 3, 3]
b_y [341, 5, 19, 96, 450, 68, 269, 0, 1269, 893, 329, 384, 932, 287, 45, 1, 17, 247, 71, 131, 812, 191, 94, 0, 199, 677, 97, 72, 7, 77, 134, 1, 1042
, 791, 187, 663, 663, 1770, 836, 0, 155, 1022, 2057, 16, 145, 249, 1365, 1, 10, 59, 110, 264, 337, 474, 248, 0, 275, 1099, 9, 225, 298, 99, 398, 1, 46, 906
, 962, 348, 69, 455, 44, 0, 750, 1966, 11, 6, 136, 39, 430, 1, 3, 3]
*****
epoch 29 batch number 347 loss is: 6.684933662414551
```

## 生成诗歌

inital linear weight

日月子邻能手语，黄柳长道岁未送，千家处梦多去日，门处乐送手人当。

inital linear weight

红书岁人孙神为，两心贫去阳中逢，更黄更一人纷作，未将明秋二神公。

inital linear weight

山水郎远行子问，花雁处来后身作，自为龙杨天外阳，柳岁四春田事莫。

inital linear weight

夜长同更归自将，人处阳色若当岁，复有春口心事有，李明心年道乐如。

inital linear weight

湖江北寻自唯四，柳一秋二多惟去，为仙径姜缀中一，水复后去足昔马。

inital linear weight

湖门事处绝和条，唯一岁秋门木雁，黄莫飞后难胜难，龙三黄龙家之寻。

inital linear weight

湖处色四清北子，水中行台田语岁，复无待三身思欢，千不欢家书阳仙。

inital linear weight

君之树阳郎莫俦，天子是五湖发生，庭行仙成惟归难，龙烟黄自落马在。