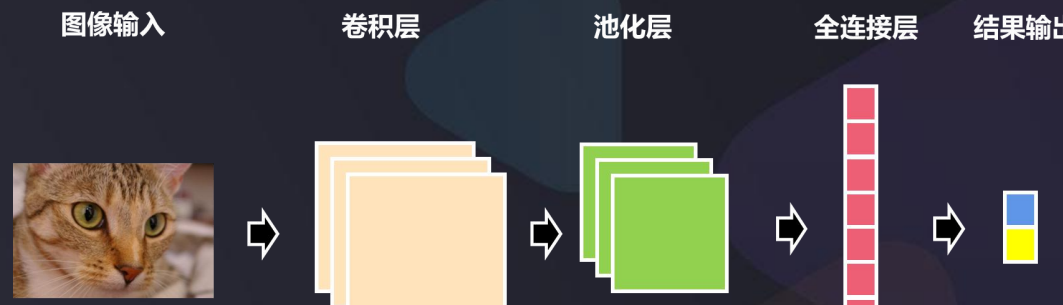
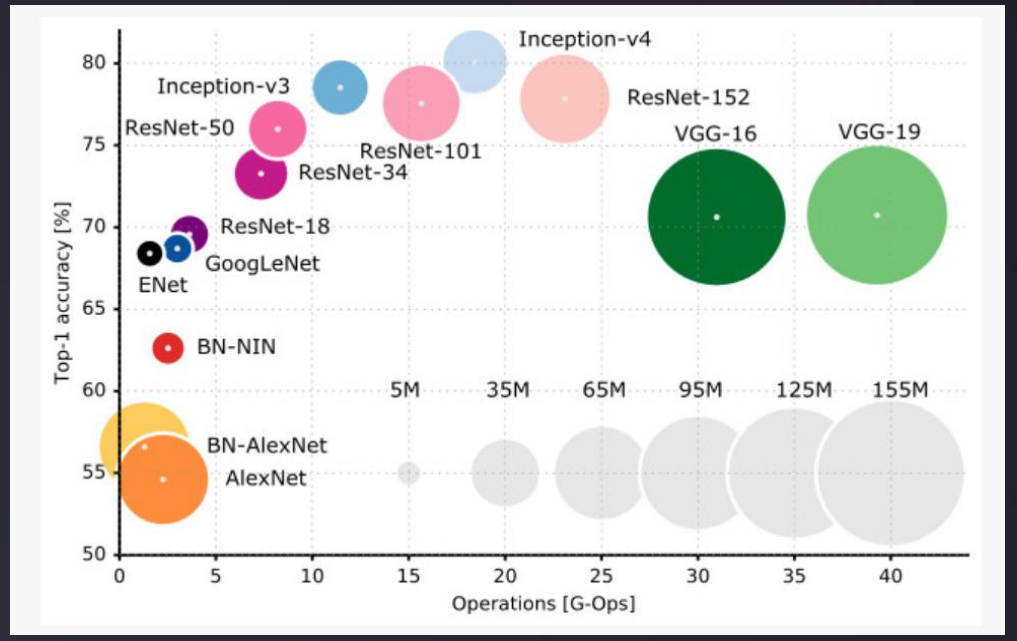
EDU
CSDN学院 IT实战派

第五章：录用之道-扎实的理论基础4

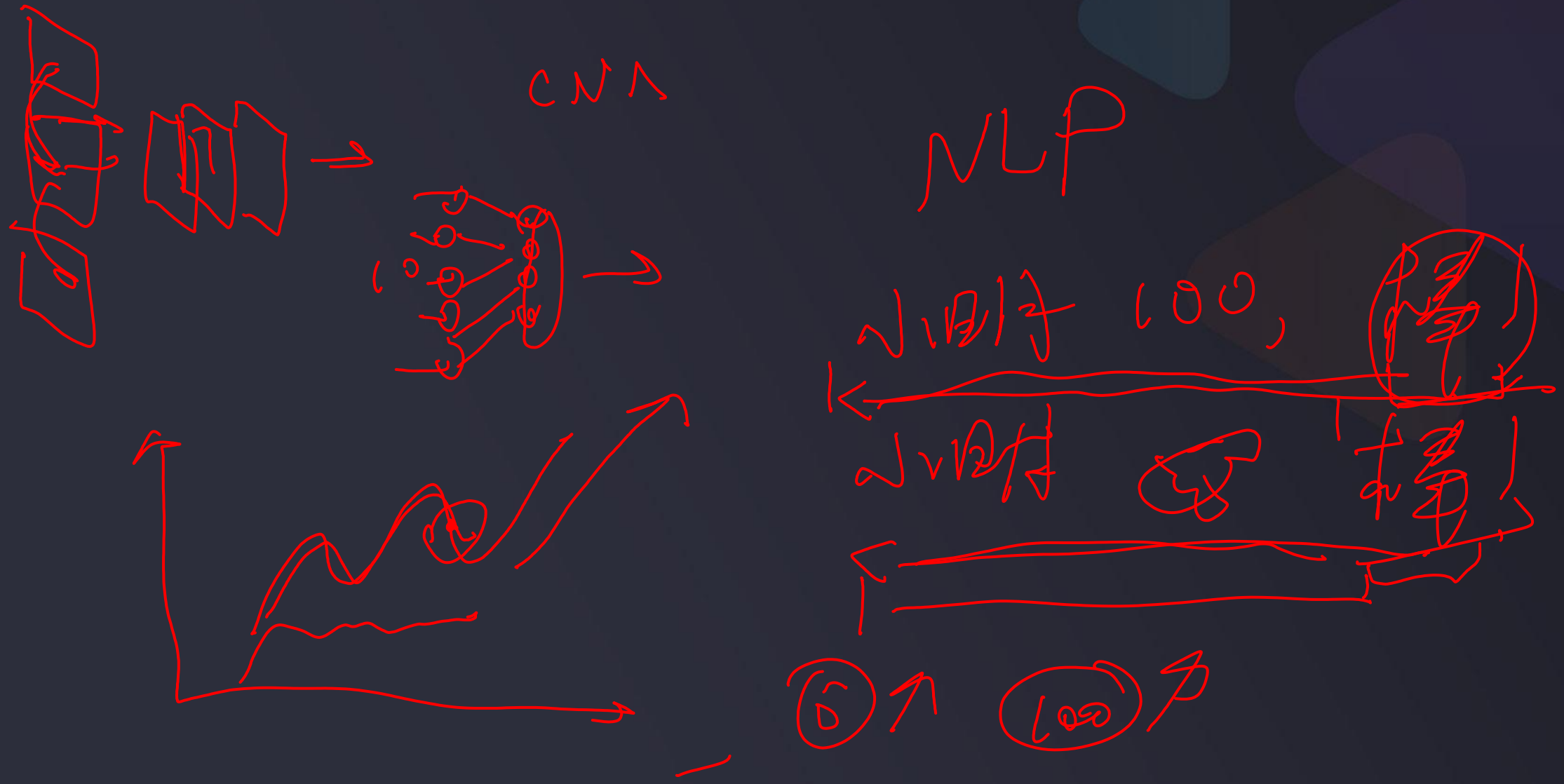
上一章回顾

- 卷积神经网络历史和背景知识
- 卷积神经网络的底层机理
- 经典的CNN结构
- 模拟面试





深度学习模型机理-循环神经网络历史和背景



深度学习模型机理-循环神经网络历史和背景

CNN
RNN

中国：我们这边快完了。
欧洲：我们这边快完了。
中国：我们好多了。
欧洲：我们好多了。

深度学习模型机理-循环神经网络历史和背景

1 语言不改变，元组恒定

2 强制字/词

3 新词

4 词向量

Recurrent Neural Network

深度学习模型机理-循环神经网络历史和背景

RNN 80-90

1982 Jhnhopfield

1986 Jordan network

1990 Jeffery Elman

BP RNN

深度学习模型机理-循环神经网络历史和背景

1997 Jürgen Schmidhuber LSTM

1997 Mike Schuster Bidirectional RNN

2014 Bengio Seq2Seq attention Mechanism

2018 BERT NLP

深度学习模型机理-循环神经网络的重要应用领域

1.语言建模和文本生成

2.机器翻译

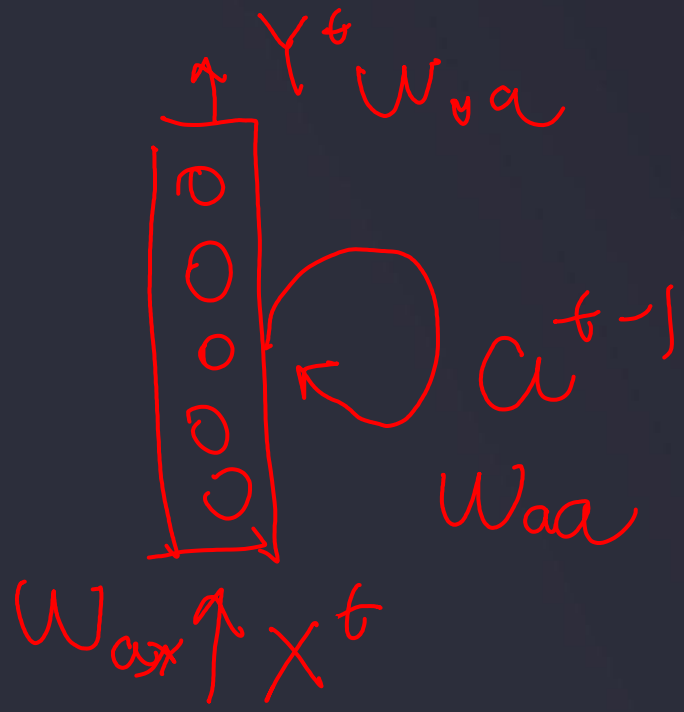
3.语音识别

4.生成图像描述

5.视频标记

CNN-RNN

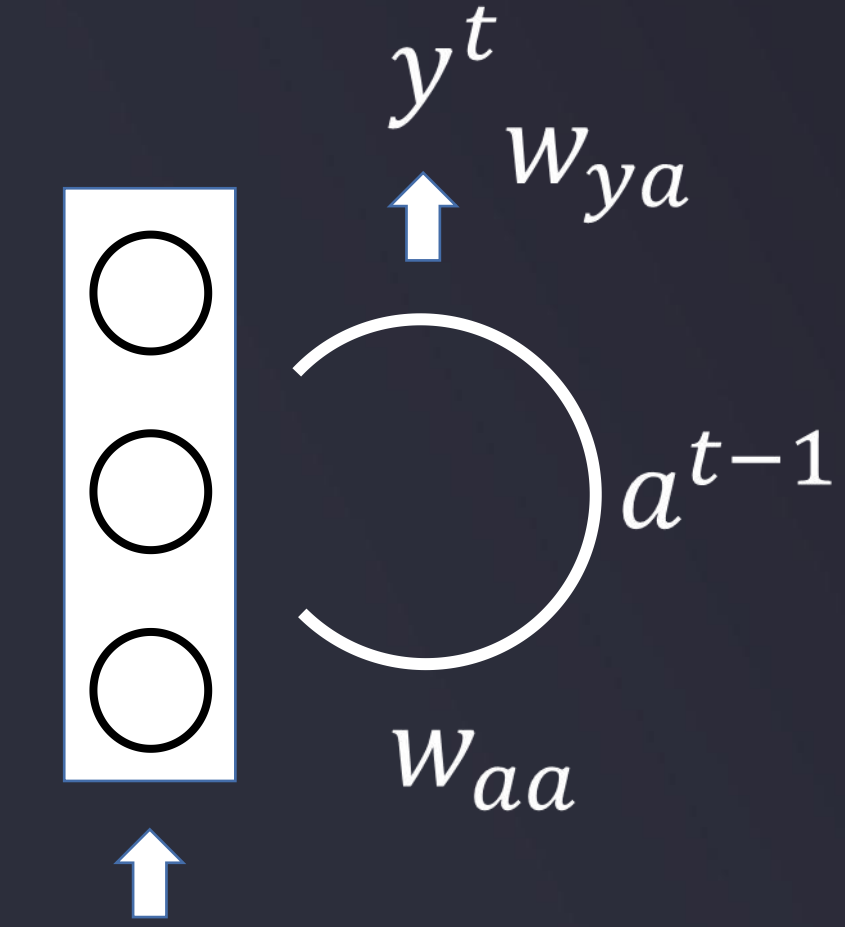
深度学习模型机理-循环神经网络的机理



$$a^t = f(W_{ax}X_t + W_{aa}a^{t-1} + b_a)$$

$$Y^t = f(W_{ya}a^t + b_y)$$

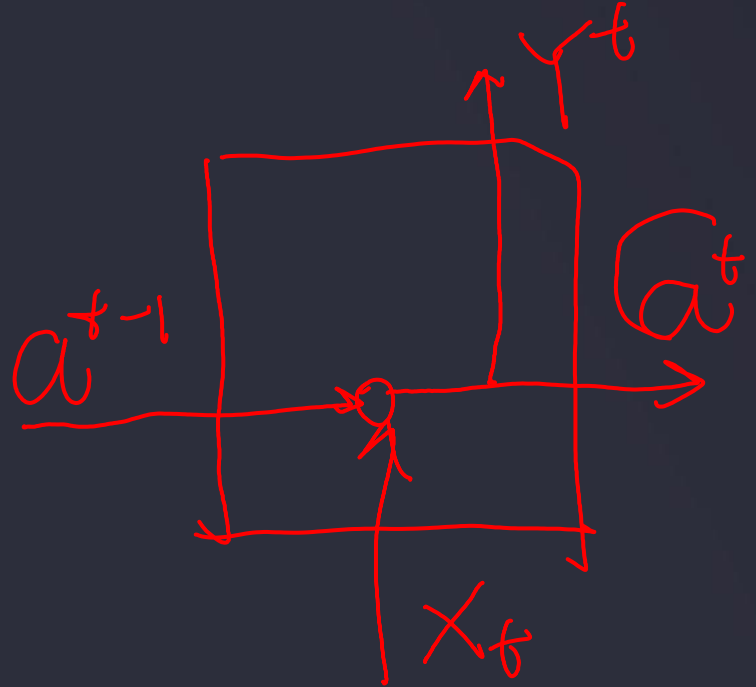
深度学习模型机理-循环神经网络的机理



$$a^t = \mathrm{f}(w_{ax}\ x^t + w_{aa}\ a^{t-1} + b_a)$$

$$y^t = \mathrm{f}(w_{ya}\ a^t + b_y)$$

深度学习模型机理-经典的循环神经网络结构GRU



a^{t+n} n ↑

a^t

Gated
Recurrent
Unit

小明 ~~~~~~, 他 ~~~~~~

小红 ~~~~~~, 她 ~~~~~~

深度学习模型机理-经典的循环神经网络结构GRU

$$c^t = f(w_c[a^{t-1}, x^t] + b_c)$$

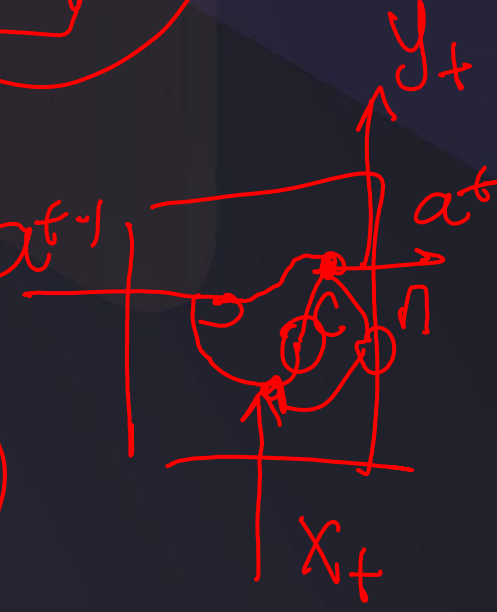
$$\Gamma_u = \sigma(w_u[a^{t-1}, x^t] + b_u)$$

sigmoid

(0-1)

$$a_t = \Gamma_u \odot c^t + (1-\Gamma_u)\, a_{t-1}$$

$$y = f(w_{ya} a_t + b_y)$$



LSTM

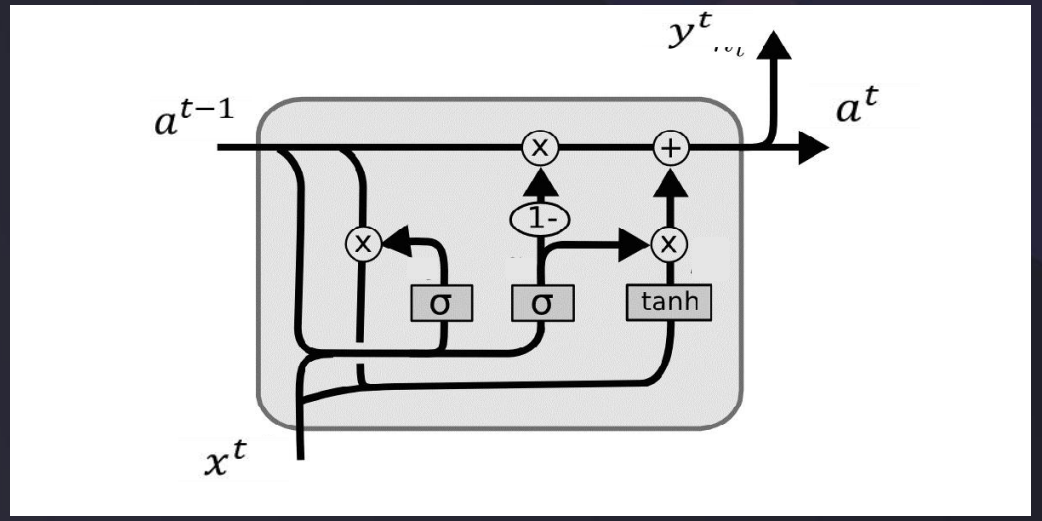
深度学习模型机理-经典的循环神经网络结构GRU

$$C^t = \mathrm{f}(w[\, x^t + a^{t-1}] + b_c)$$

$$\Gamma_u = \mathrm{б}(w[\, x^t + a^{t-1}] + b_u)$$

$$a^t = \Gamma_u C^t + (1 - \Gamma_u)\, a^{t-1}$$

$$y^t = \mathrm{f}(w_{ya}\, a^t + b_y)$$



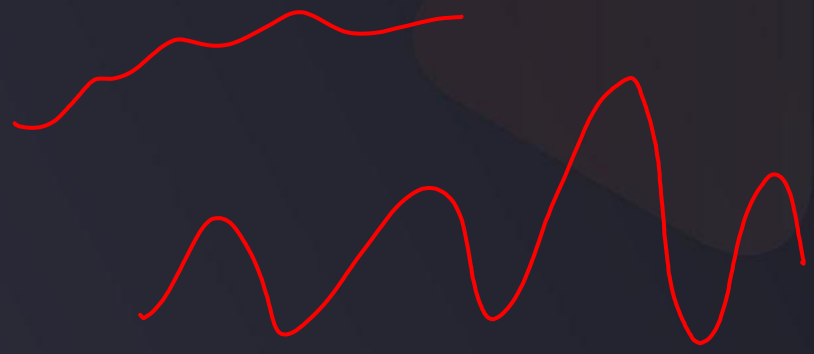
LSTM

深度学习模型机理-经典的循环神经网络结构LSTM

Long Short Term Memory

1997

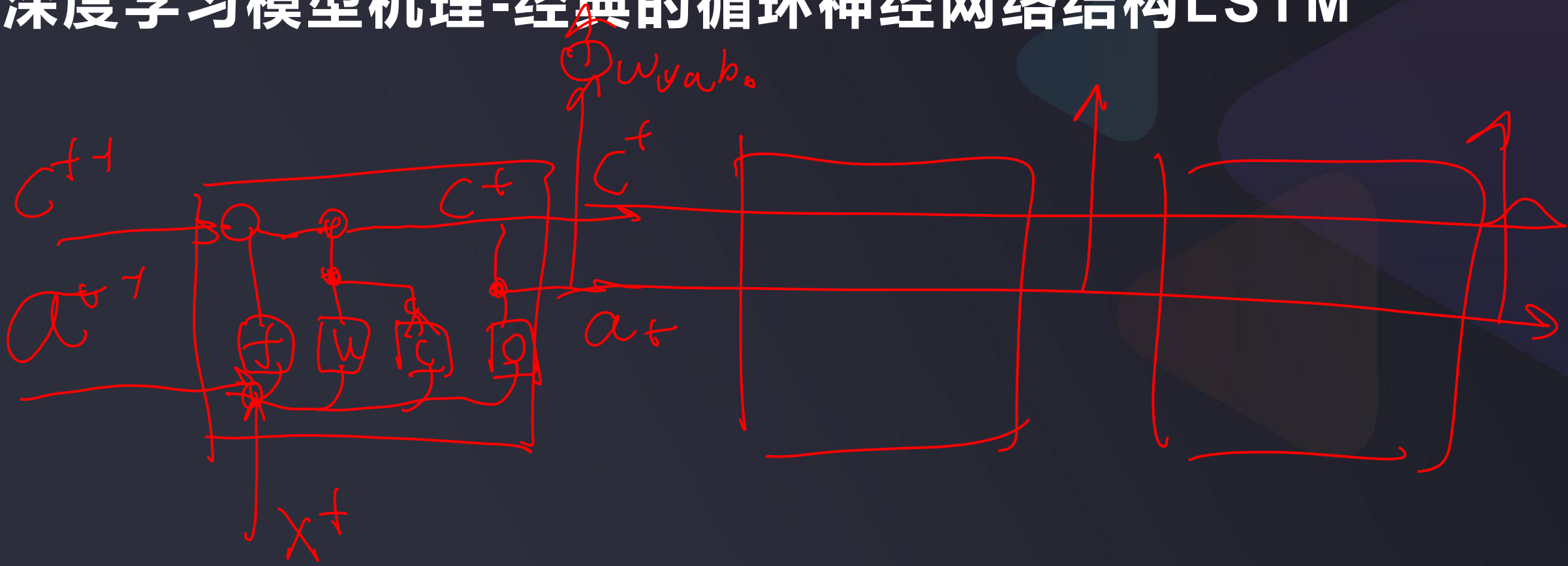
2009 ICDAR



深度学习模型机理-经典的循环神经网络结构LSTM

$$
\begin{cases}
\tilde{c}_t = \tanh\left(W_c\left[a^{t-1}, x^t\right] + b_c\right) \\
\Gamma_u = \sigma\left(W_u\left[a^{t-1}, x^t\right] + b_u\right) \\
\Gamma_f = \sigma\left(W_f\left[a^{t-1}, x^t\right] + b_f\right) \\
\Gamma_o = \sigma\left(W_o\left[a^{t-1}, x^t\right] + b_o\right) \\
c^t = \Gamma_u * \tilde{c}_t + \Gamma_f * c_{t-1} \\
a^t = \Gamma_o * c^t
\end{cases}
$$

深度学习模型机理-经典的循环神经网络结构LSTM



深度学习模型机理-经典的循环神经网络结构LSTM

$$\hat{C}^t = \tanh(w_c[x^t + a^{t-1}] + b_c)$$

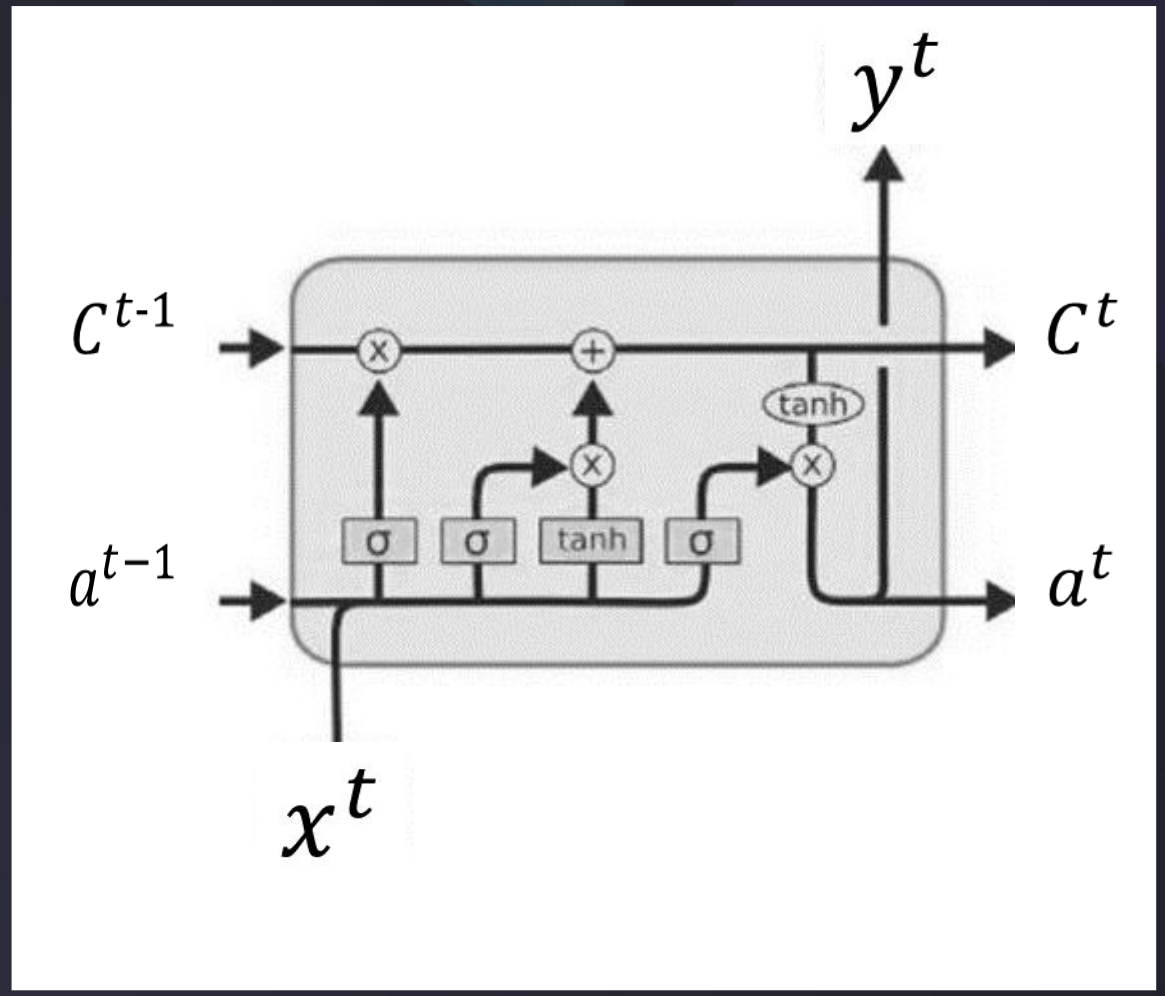
$$\Gamma_u = \sigma(w_u[x^t + a^{t-1}] + b_u)$$

$$\Gamma_f = \sigma(w_f[x^t + a^{t-1}] + b_f)$$

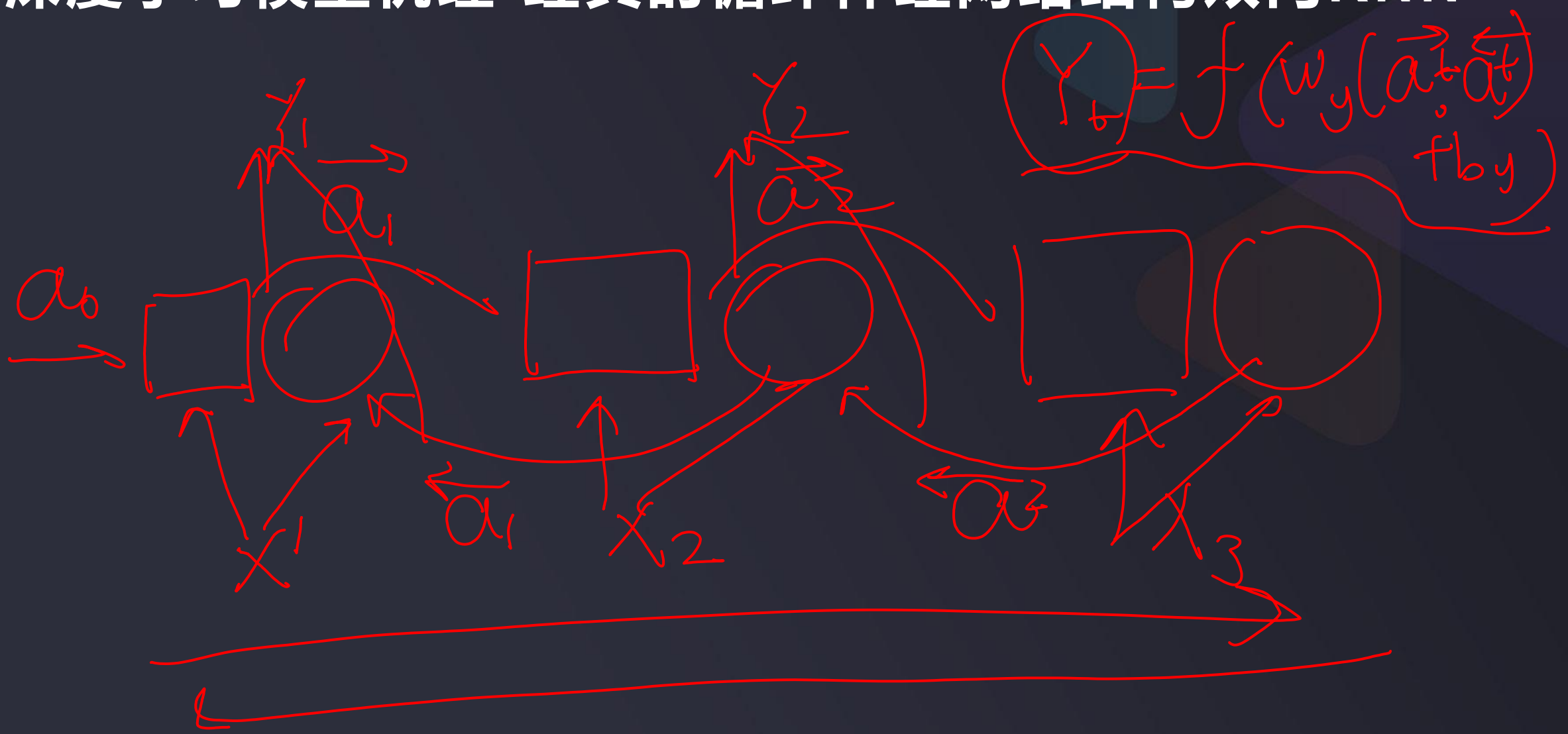
$$\Gamma_o = \sigma(w_o[x^t + a^{t-1}] + b_o)$$

$$C^t = \Gamma_u \hat{C}^t + \Gamma_f C^{t-1}$$

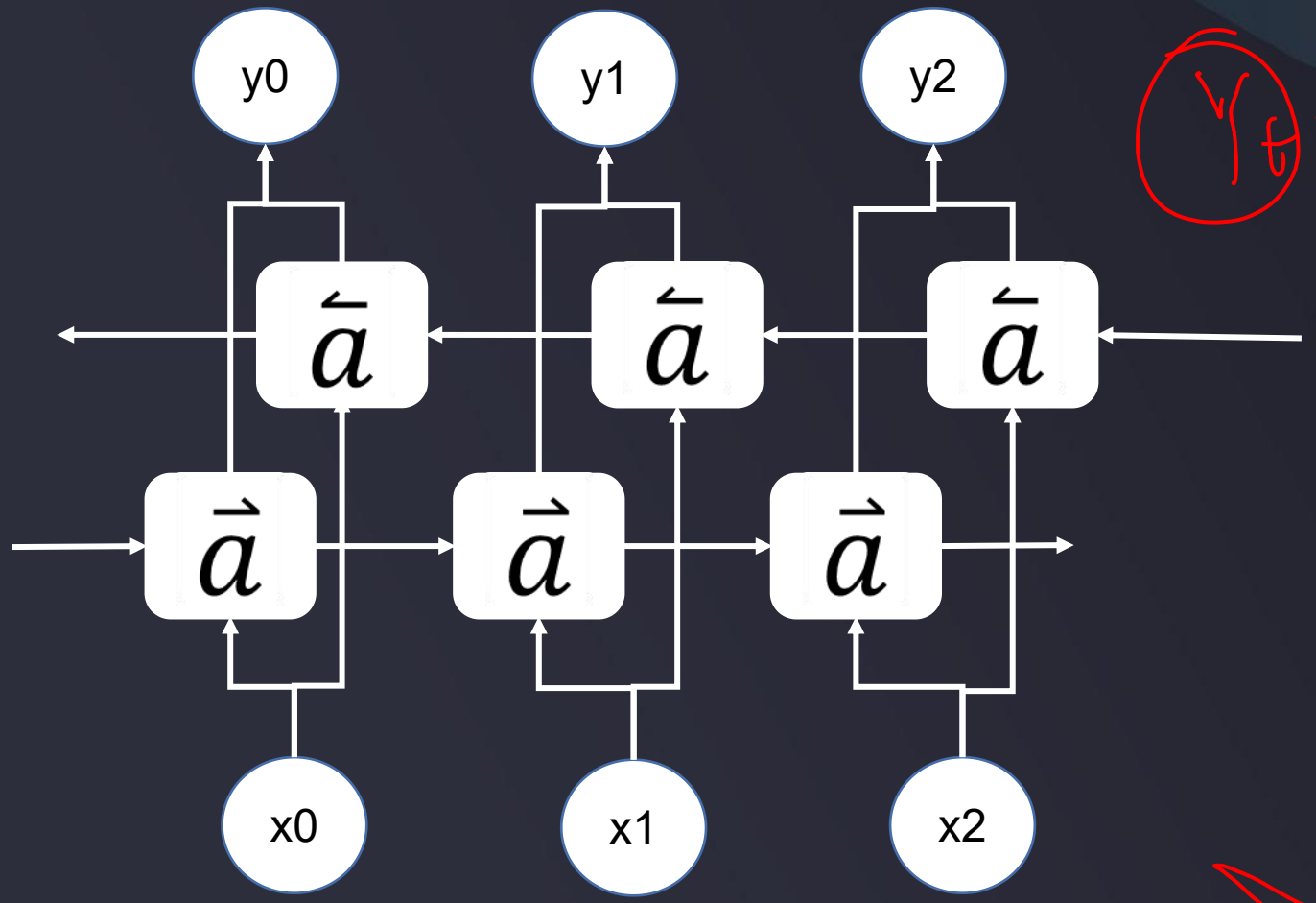
$$a^t = \Gamma_o C^t$$



深度学习模型机理-经典的循环神经网络结构双向RNN

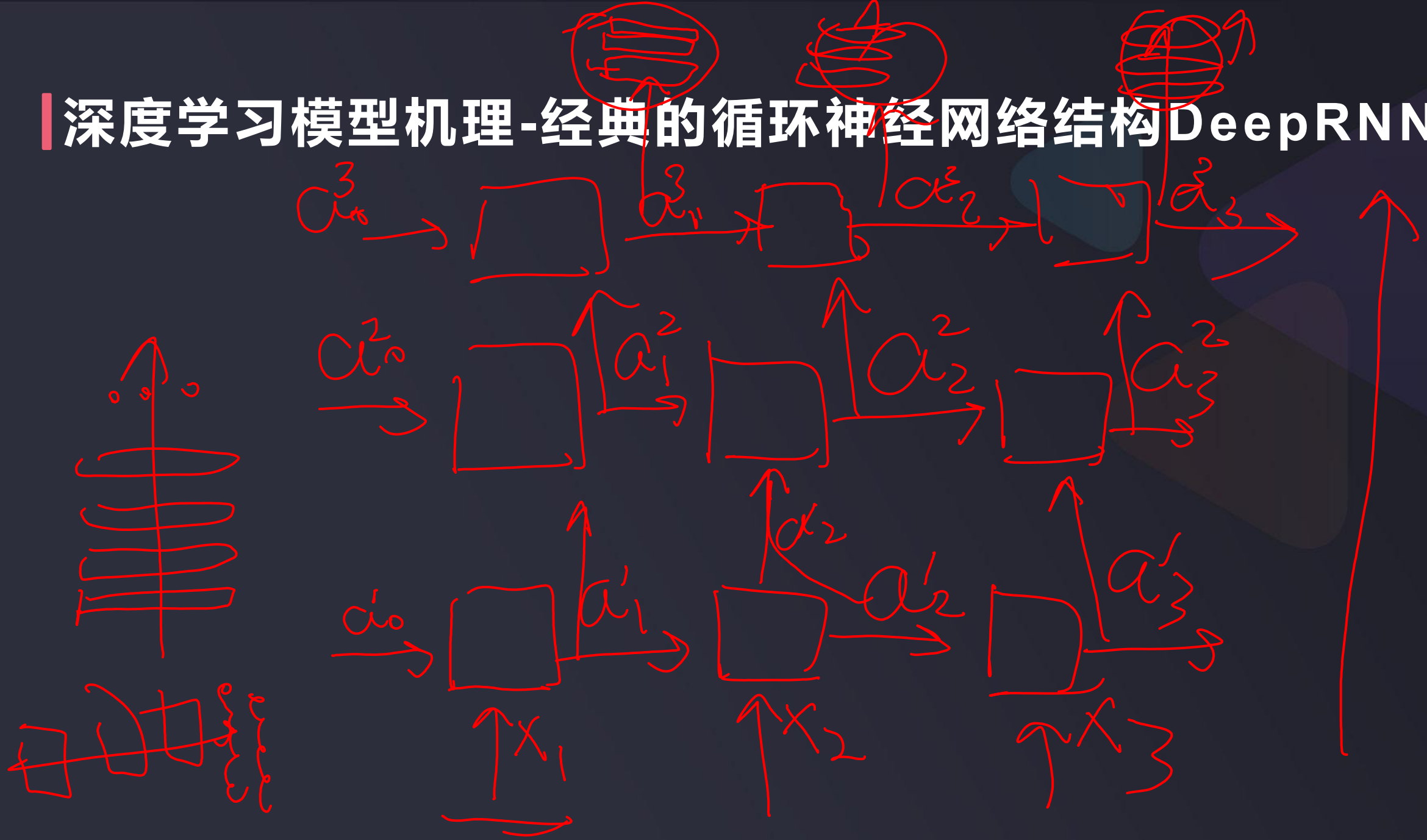


深度学习模型机理-经典的循环神经网络结构双向RNN

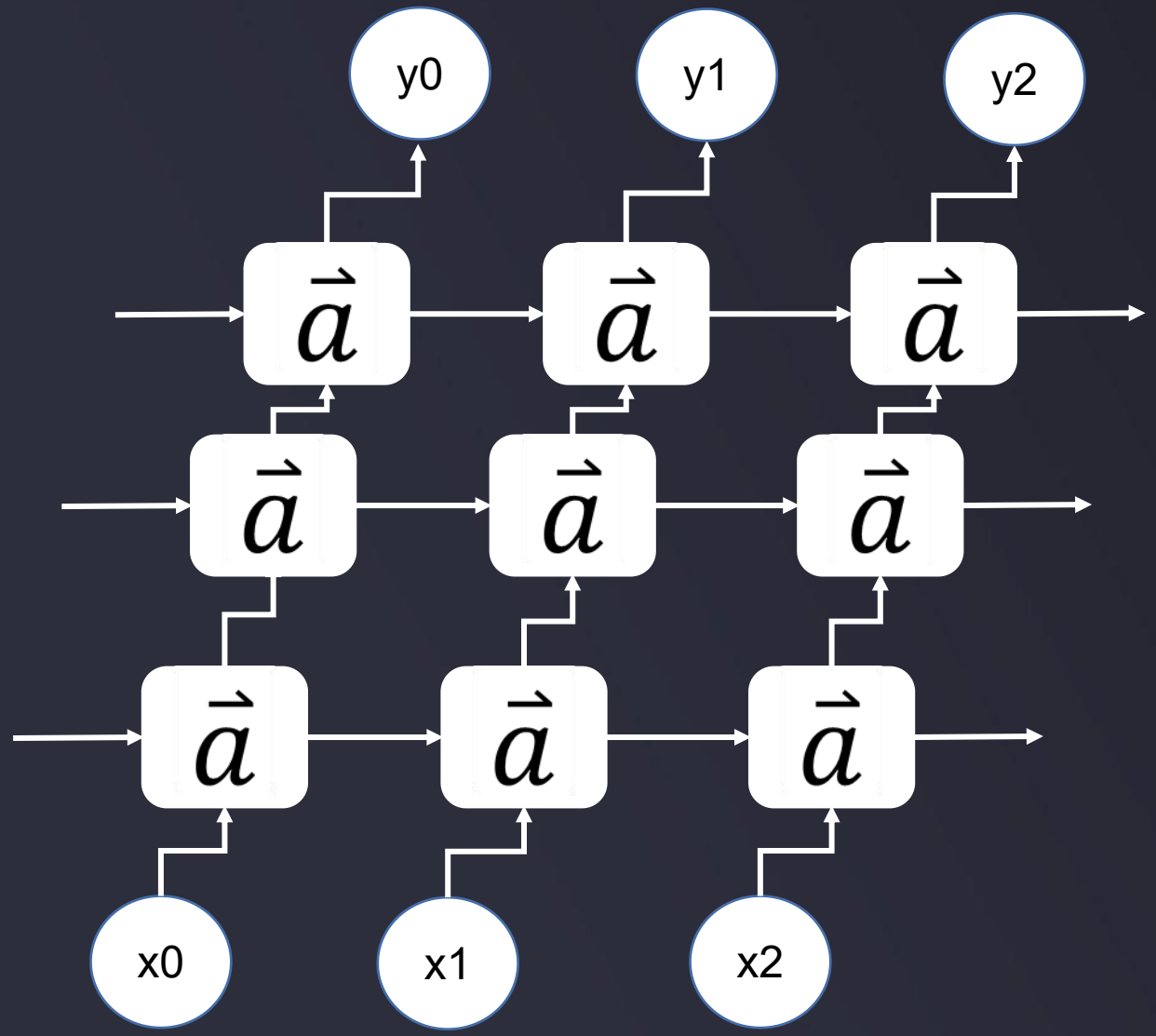


$$y_t = f(w_y(\vec{a}^t, \overleftarrow{a}^t) + b_y)$$

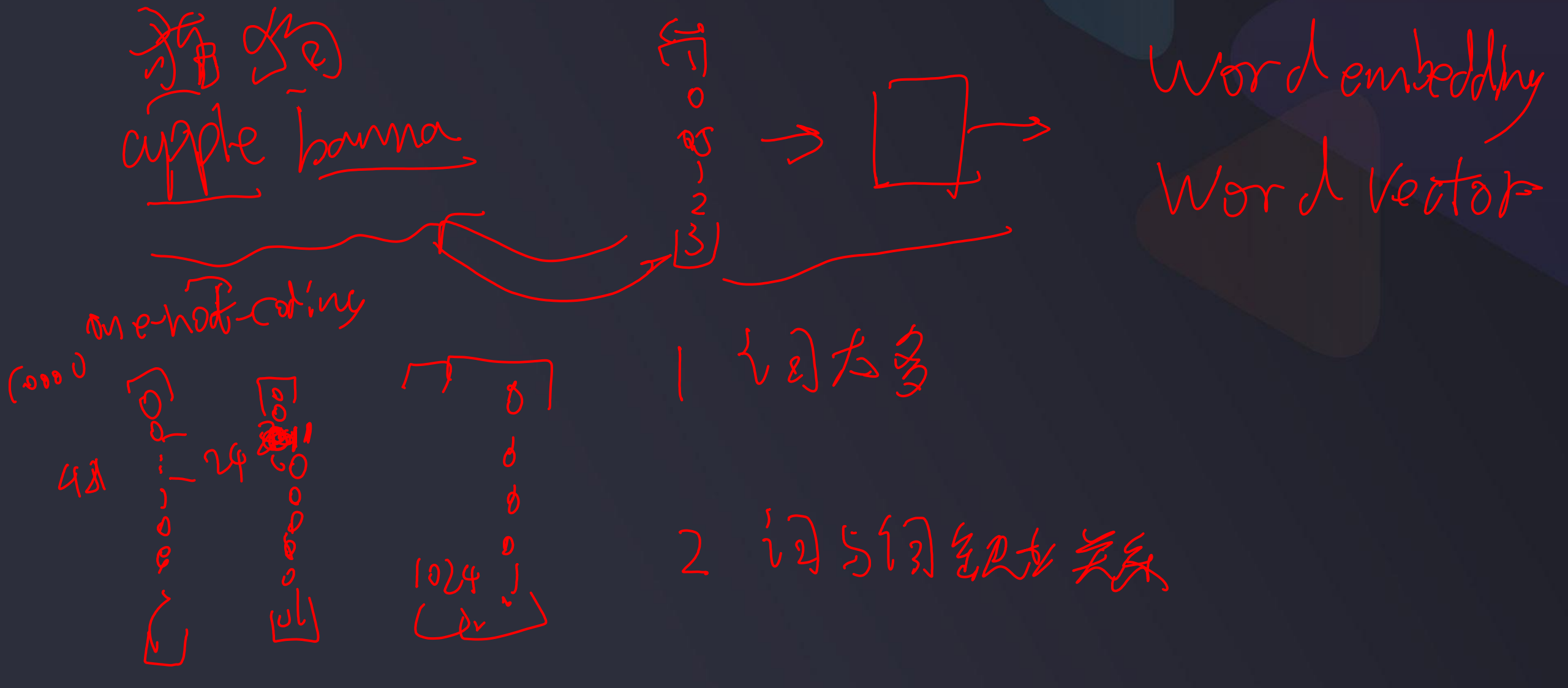
深度学习模型机理-经典的循环神经网络结构DeepRNN



深度学习模型机理-经典的循环神经网络结构DeepRNN



深度学习模型机理-自然语言处理词汇表征和词嵌入向量



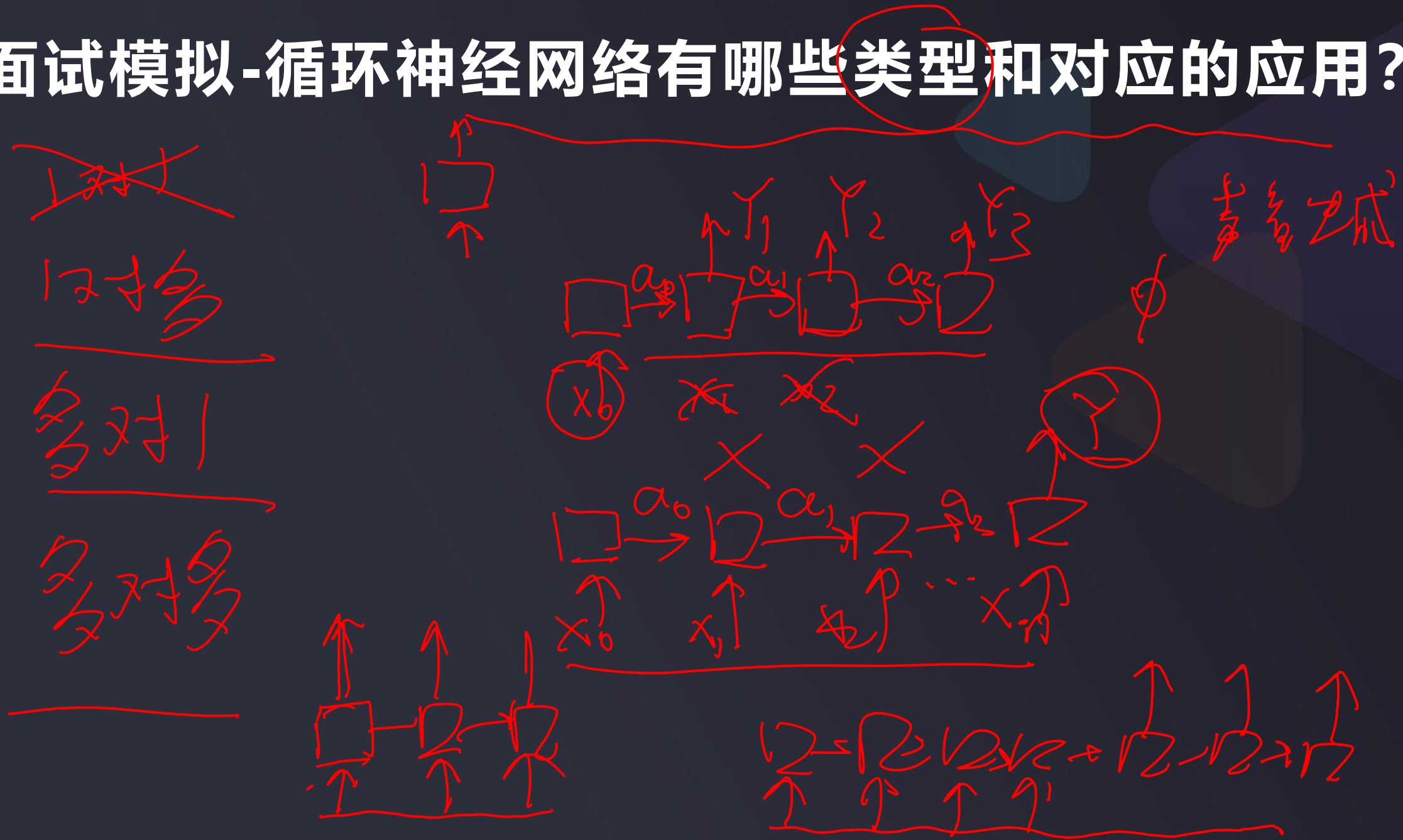
深度学习模型机理-自然语言处理词汇表征和词嵌入向量

	苹果	橙子	苹果	手机
动物	0.99	0.97	0.1	0.05
水果	0.05	0.02	0.99	0.99
红色	0.2	0.1	0.8	0.3
发声音	0.9	0.95	0.05	0.12

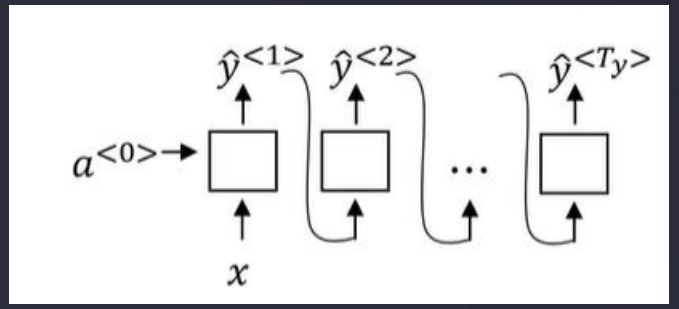
深度学习模型机理-自然语言处理词汇表征和词嵌入向量

Keras

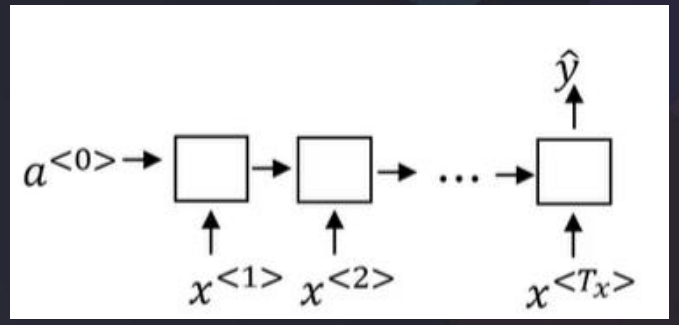
面试模拟-循环神经网络有哪些类型和对应的应用？



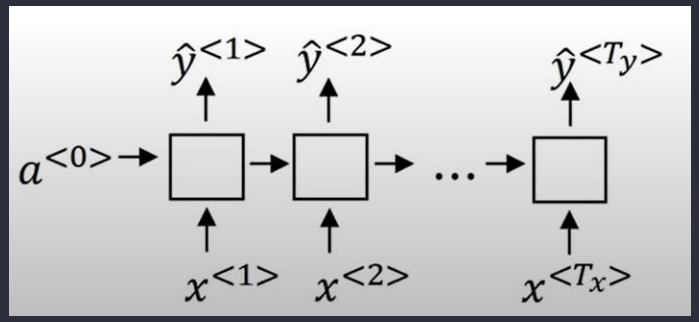
面试模拟-循环神经网络有哪些类型和对应的应用？



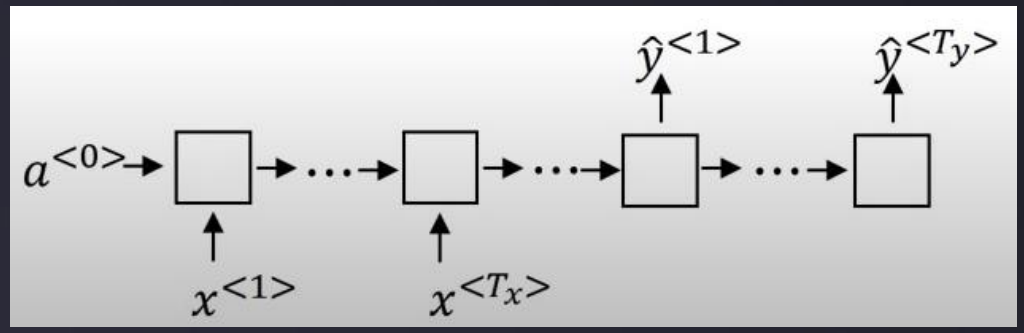
1对多
应用例子：音乐/诗歌生成



多对1
应用例子：情绪识别，文档分类



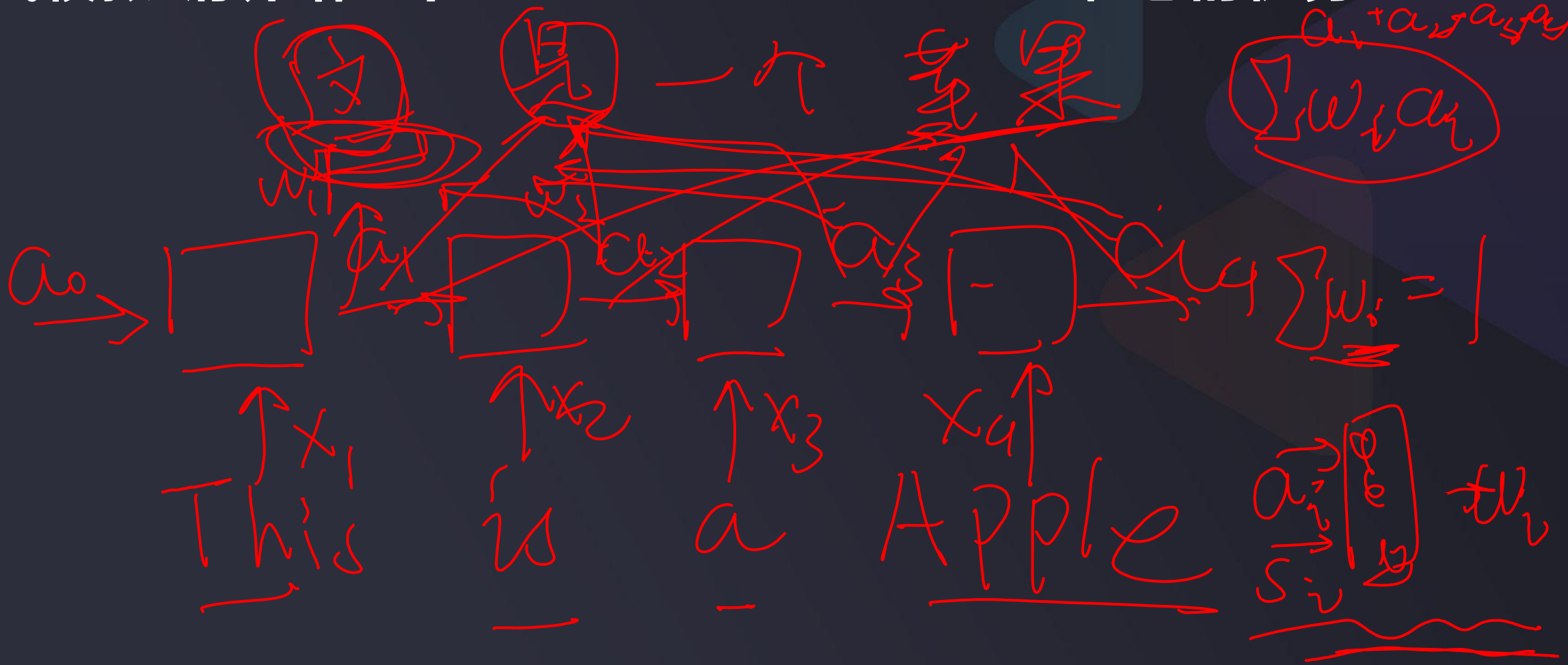
多对多，类型一
应用例子：金融预测，时间序列预测



多对多，类型二
应用例子：机器翻译

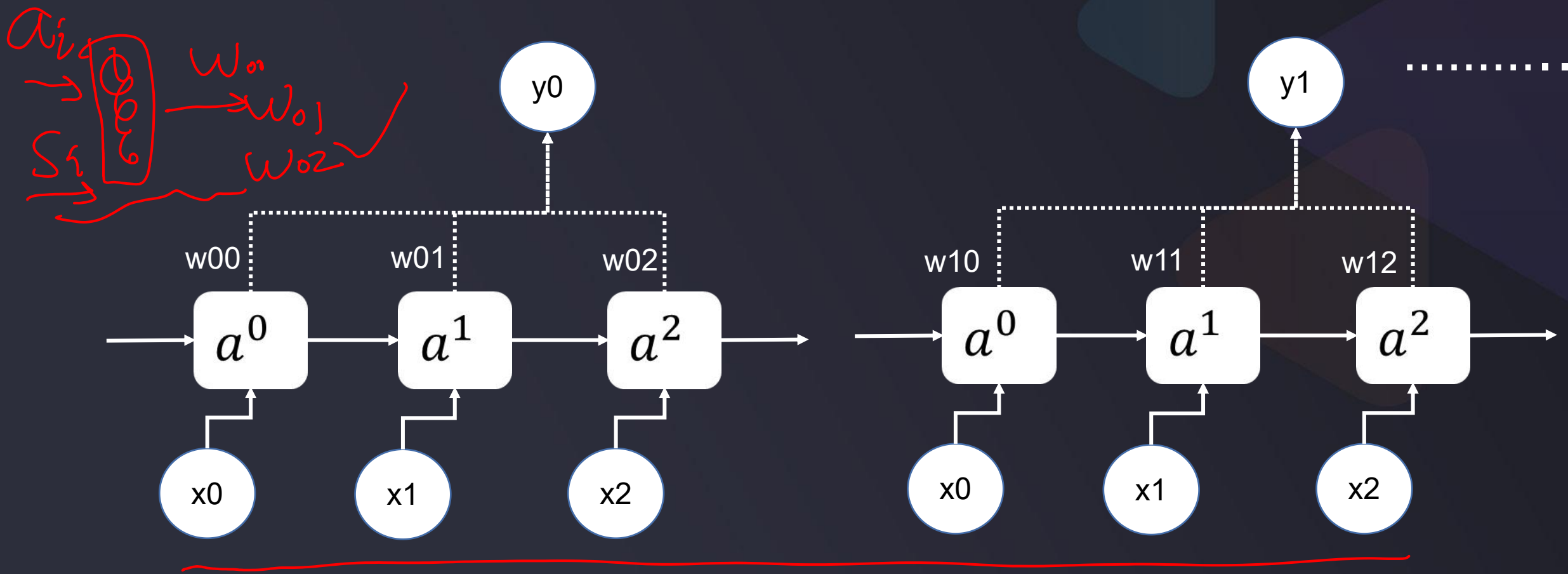
面试模拟-请介绍一下attention mechanism和它的优势?

面试模拟-请介绍一下attention mechanism和它的优势?



Mnih V, Heess N, Graves A. Recurrent models of visual attention[C]//Advances in Neural Information Processing Systems. 2014: 2204-2212.

面试模拟-请介绍一下attention mechanism和它的优势?



Mnih V, Heess N, Graves A. Recurrent models of visual attention[C]//Advances in Neural Information Processing Systems. 2014: 2204-2212.

本章回顾

- 循环神经网络历史和背景
- 循环神经网络的重要应用领域
- 循环神经网络的机理
- 循环神经网络的经典结构
- 自然语言处理词汇表征和词嵌入向量
- 面试模拟-循环神经网络的类型和注意力机制

课程相关资料

欢迎大家扫码或者添加微信好友ai_flare（学习小助手），加入学习群，老师会在群里帮大家解答学习、职业发展与求职问题（名额有限、人满即止）