

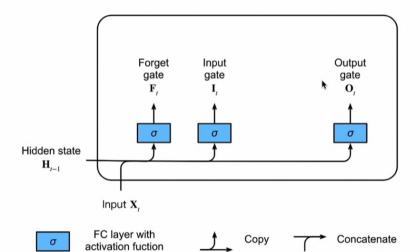
长短期记忆网络

• 忘记门:将值朝0减少

• 输入门: 决定不是忽略掉输入数据

・输出门: 决定是不是使用隐状态

$$\begin{split} \boldsymbol{I}_t &= \sigma(\boldsymbol{X}_t \boldsymbol{W}_{xi} + \boldsymbol{H}_{t-1} \boldsymbol{W}_{hi} + \boldsymbol{b}_i) \\ \boldsymbol{F}_t &= \sigma(\boldsymbol{X}_t \boldsymbol{W}_{xf} + \boldsymbol{H}_{t-1} \boldsymbol{W}_{hf} + \boldsymbol{b}_f) \\ \boldsymbol{O}_t &= \sigma(\boldsymbol{X}_t \boldsymbol{W}_{xo} + \boldsymbol{H}_{t-1} \boldsymbol{W}_{ho} + \boldsymbol{b}_o) \end{split}$$



动手学深度学习 v2・<u>https://courses.d2l.ai/zh-v2</u>



• 忘记门:将值朝0减少

・输入门: 决定不是忽

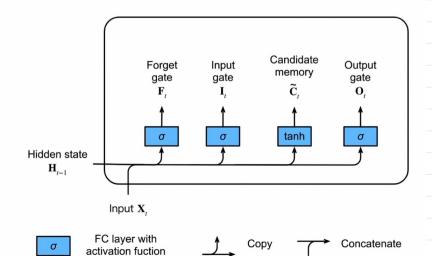
略掉输入数据

・输出门: 决定是不是

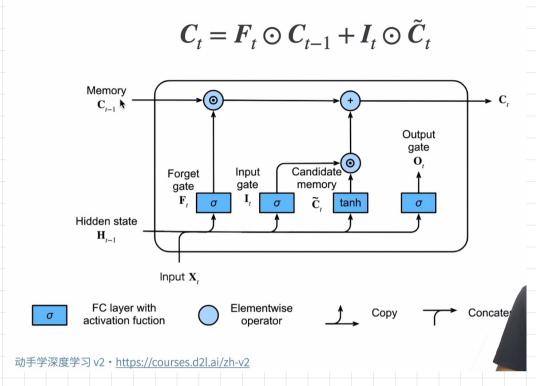


候选记忆单元

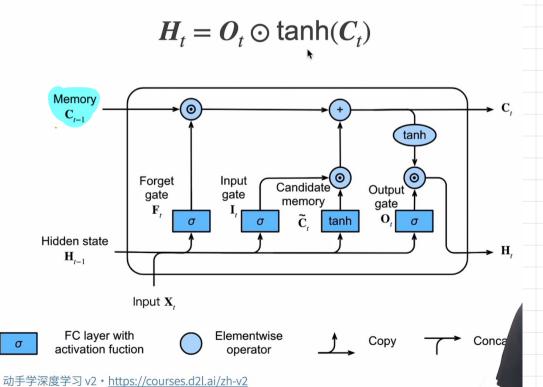
$$\tilde{\boldsymbol{C}}_t = \tanh(\boldsymbol{X}_t \boldsymbol{W}_{xc} + \boldsymbol{H}_{t-1} \boldsymbol{W}_{hc} + \boldsymbol{b}_c)$$



记忆单元

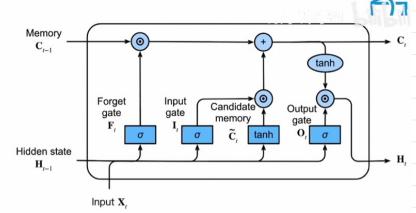


隐状态



总结

$$egin{aligned} oldsymbol{I}_t &= \sigma(X_t W_{xi} + oldsymbol{H}_{t-1} W_{hi} + oldsymbol{b}_i) \ oldsymbol{F}_t &= \sigma(X_t W_{xf} + oldsymbol{H}_{t-1} W_{hf} + oldsymbol{b}_f) \ oldsymbol{O}_t &= \sigma(X_t W_{xo} + oldsymbol{H}_{t-1} W_{ho} + oldsymbol{b}_o) \ oldsymbol{ ilde{C}}_t &= anh(X_t W_{xc} + oldsymbol{H}_{t-1} W_{hc} + oldsymbol{b}_c) \ oldsymbol{C}_t &= oldsymbol{F}_t \odot oldsymbol{C}_{t-1} + oldsymbol{I}_t \odot oldsymbol{ ilde{C}}_t \ oldsymbol{H}_t &= oldsymbol{O}_t \odot anh(oldsymbol{C}_t) \end{aligned}$$





动手学深度学习 v2·https://courses.d2l.ai/zh-v2