

## LSTM based



```
graph LR; A[LSTM based] --- B[梯度更新公式a'=a-learning_rate*grad, 可以类比LSTM的记忆门的更新, 参数a即为记忆变量]; A --- C[将梯度输入LSTM更新作为记忆量的参数]
```

梯度更新公式 $a' = a - \text{learning\_rate} * \text{grad}$ , 可以类比LSTM的记忆门的更新, 参数 $a$ 即为记忆变量

将梯度输入LSTM更新作为记忆量的参数