

به نام خدا

امیر محمد عقیبان زاده

۴۰۱۳۰۰۵۲۵

تیمین و یارگیری عمیق

استاد: دکتر فاطمی زاده

۱) در هر بازه به رابطه چهارم اگر z_t صفر باشد، حافظه قبلی (h_t) در نظر گرفته نمی شود و مؤثر آن اطلاعات ورودی جدید اثر ندارد و فراموش می شود. z_t کنترل کننده میزان ارزش دهی به حالت قبلی است و در واقع یکیت $update$ است.
با توجه به رابطه سوم اگر z_t صفر باشد، اثر حالت قبلی از میان می رود. با مشخص شدن z_t به چه میزان در دو حالت قبلی، مقدار حالت جدید را تعیین می کند و در واقع یکیت $reset$ است.

ب) ۱

$$\frac{\partial L}{\partial w} = \sum_{j=0}^{T-1} \sum_{k=1}^J \frac{\partial L_j}{\partial y_j} \frac{\partial y_j}{\partial h_j} \left(\prod_{m=k+1}^J \frac{\partial h_m}{\partial h_{m-1}} \right) \frac{\partial h_k}{\partial w}$$

$$\frac{\partial h_m}{\partial h_{m-1}} \xrightarrow{RNN} \frac{\partial h_t}{\partial h_{t-1}} = [1 - \tanh^2(w_{hh} h_{t-1} + w_{hz} z_t)] w_{hh}$$

$$\frac{\partial h_t}{\partial h_{t-1}} \xrightarrow{XNN} \frac{\partial h_t}{\partial h_{t-1}} = z_t + (1 - z_t) \frac{\partial \tilde{h}_t}{\partial h_{t-1}} \rightarrow \text{اگر } z_t \text{ برابر ۱ باشد در ضرب با هم متوالی رابطه}$$

$\frac{\partial L}{\partial w} = \text{grad. vanishing}$ یعنی فراموشی داشت.