

تمرین چهارم - کدایش فزونی کدوری ۹۸۱۰۲۵۴۲

۱) الانم با توجه به رابطه آخر که برای بهذرسی حالت شبکه است، اگر z_t برابر صفر باشد، حلقه قبل در تفرقه نمی شود و
موترا اطلاعات ورودی جدید اثر گذار خواهد بود. z_t کنترل کننده میزان ارزش دهی به حالت قبل است، مانند کت update
محلی می کند.

از طرفی اگر z_t برابر صفر باشد، اثر حالت قبل مطلقاً از بین می رود. z_t مشخص می کند به چه میزان ورودی و یا حالت قبل
مقدار حالت جدید را تعیین می کند. z_t هتند کت Reset محلی می کند.

$$\frac{\partial L}{\partial W} = \sum_{j=0}^{T-1} \sum_{k=0}^j \frac{\partial L_j}{\partial y_j} \cdot \frac{\partial y_j}{\partial h_j} \cdot \frac{\partial h_j}{\partial h_k} \cdot \frac{\partial h_k}{\partial W} \quad (ب)$$

می دانیم در Vanilla RNN این حلقه حاصل

فقدان های مقدار است که سبب Gradient Vanish
هست.

$$XNN: \frac{\partial h_j}{\partial h_k} = \prod_{m=k+1}^j \frac{\partial h_m}{\partial h_{m-1}} \rightarrow \frac{\partial h_m}{\partial h_{m-1}} = z_m + (1-z_m) \frac{\partial \tilde{h}_m}{\partial h_{m-1}}$$

$$z_m = 1 \Rightarrow \frac{\partial h_m}{\partial h_{m-1}} = 1 \Rightarrow \text{No Gradient Vanishing!}$$

در اینجا مهم ترین عامل مشکل از تبدیل به ۱ می شود.