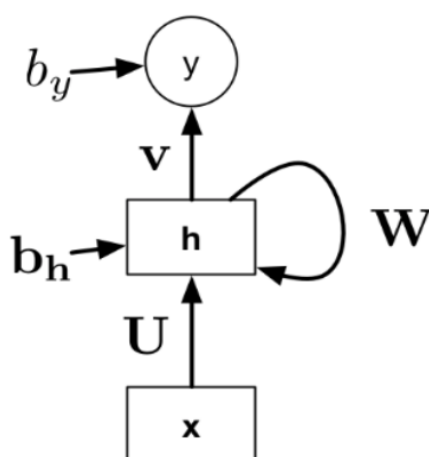




مسئله‌ی ۱. جمع‌کننده دودویی

مسئله جمع دودویی (binary) را در نظر بگیرید. فرض کنید بخواهیم یک شبکه MLP را آموزش دهیم که این عملیات را انجام دهد. مشکلاتی که با آن مواجه خواهیم شد چه خواهد بود؟

حال یک شبکه بازگشتی پیشنهاد دهید که این مسئله را حل کند. در step زمانی t ، یک ورودی دو بعدی binary داریم که هر یک نشان‌دهنده بیت t هر یک از اعداد می‌باشد. بردار h را یک بردار سه بعدی در نظر بگیرید و خروجی y را یک بعدی. شما می‌بایست مقادیری برای W ، U ، b_y ، b_h ، v ارائه دهید که مسئله را حل کند. تابع فعالسازی را تابع پله در نظر بگیرید. همچنین هر یک از sequence های ورودی را در صورت نیاز می‌توانید با صفر extend کنید.



مسئله‌ی ۲. ترکیبی

- دو مورد از مشکلات استفاده از بازنمایی one-hot برای نمایش کلمات را توضیح دهید. RNN چگونه این مشکلات را حل می‌کند؟
- مزیت استفاده از مکانیزم توجه نسبت به RNN چیست؟ توضیح دهید.
- مزیت اصلی‌ای که transformer در برابر LSTM برای ما ایجاد می‌کند چیست؟ به طور خلاصه توضیح دهید چگونه این کار را انجام می‌دهد.

- در بخش decoder مدل Transformer ، مکانیزم توجهی که روی ورودی decoder اعمال می‌شود Masked Attention نام دارد که یعنی هر کلمه ورودی در decoder ، برای محاسبه attention تنها به کلمه‌های قبلی در ورودی دسترسی دارد و کلمات بعدی به اصطلاح mask می‌شوند تا کلمه‌های قبل به آن‌ها دسترسی نداشته باشند. علت این موضوع را شرح دهید.

موفق باشید (:)