

به نام خدا

در معماری های قبلی encoder معنی جمله را به یک hidden state با طول ثابت تبدیل میکند که انتظار می رود این بازنمایی یک خلاصه خوب از تمام جمله ورودی باشد و سپس این بردار که آخرین state از encoder میباشد را به به decoder میدهد تا ترجمه متناظر را تولید کند. اشکال واضح این معماری همین hidden state با طول ثابت است که قابلیت به خاطر سپردن جملات بلند را ندارد و قسمت های اولیه جمله را فراموش میکند.

راه حل مقاله: برای پیش بینی هر کلمه مورد نظر از بیشترین کلمات مرتبط در جمله ورودی استفاده میکند تا بتواند ترجمه را بهبود ببخشد.

راه حل و به دست آوردن context vector:

$$p(y_i | y_1, \dots, y_{i-1}, \mathbf{x}) = g(y_{i-1}, s_i, c_i),$$

برای پیش بینی هر کلمه از یه context vector به اسم  $c_i$  استفاده میکنیم.

$$c_i = \sum_{j=1}^{T_x} \alpha_{ij} h_j.$$

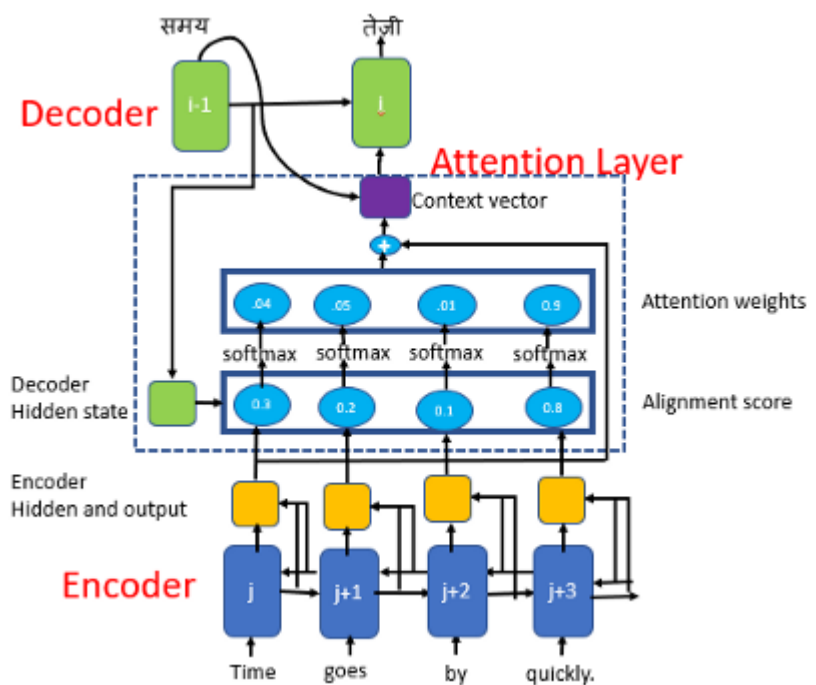
این  $c_i$  از جمع وزن دار هر کدام از hidden state های encoder (متناظر با هر کلمه ورودی) به دست می آید.

$$\alpha_{ij} = \frac{\exp(e_{ij})}{\sum_{k=1}^{T_x} \exp(e_{ik})},$$

این وزن ها از خروجی softmax متغیر e به دست می آید که به آن alignment score میگوییم.

$$e_{ij} = a(s_{i-1}, h_j)$$

این alignment score تناظر بین ورودی های اطراف j و خروجی در پوزیشن i را نگاشت میکند و اینگونه attention mechanism اعمال میشود.



معماری این روش به صورت شکل بالا میباشد.