

محمدرضا غفرانی
۴۰۰۱۳۱۰۷۶
۲۴ تیر ۱۴۰۱

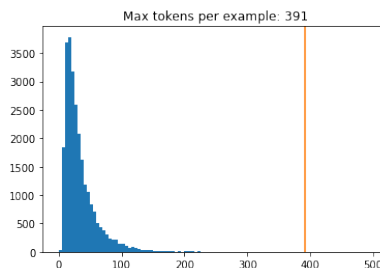
شبکه‌های عصبی

تمرین هشتم

سوال یک

برای تبدیل جملات به توکن‌ها از مدلی که برای این مجموعه داده بهینه شده است^۱ استفاده می‌کنیم. این توکن‌کننده علاوه بر تبدیل جملات به توکن‌ها پیش‌پردازش‌های دیگری نظیر یکسان‌سازی^۲ متون و کوچک‌کردن^۳ تمامی کلمات را انجام می‌دهد. پیش‌پردازش دیگری که ما انجام می‌دهیم حذف جملاتی که طول آن‌ها بسیار زیاد است. هدف از حذف این جملات به‌ترکردن عملکرد مدل با حذف این داده‌های پرت است.

برای حذف کلماتی که اندازه آن‌ها از یک حد آستانه‌ای بیشتر است، ابتدا نیاز است تا درکی از طول جملات موجود در مجموعه داده داشته باشیم. بنابراین طول تمامی جملات را در مجموعه داده استخراج و در شکل ۱ رسم می‌کنیم. بنابراین این شکل طول جملات معمولاً تا ۱۰۰ کلمه است این حالی که در همین مجموعه داده جمله‌ای با طول ۳۹۱ کلمه نیز وجود دارد. بنابراین توضیحات ۱۲۸ طول معقولی برای حداکثر جمله به نظر می‌رسد.



شکل ۱: طول جملات موجود در مجموعه داده

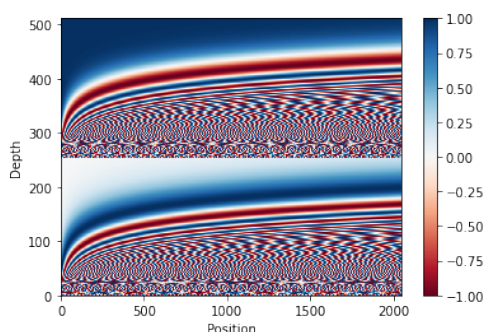
سوال دوم

در شکل ۲ تعبیه مکانی که برای هر کلمه در نظر گرفته می‌شود آورده شده است. این تعبیه مکانی برای جمله‌ای در اندازه ۲۰۴۸ کلمه قابل استفاده است. ما در این جا

^۱https://www.tensorflow.org/text/guide/subwords_tokenizer

^۲Normalize

^۳lowercase



شکل ۲: تعبیه مکانی برای هر کلمه در جمله‌ای به طول ۲۰۴۸

با توجه به این که طول جملات ورودی بسیار کوچک‌تری داریم (طول جملات ورودی و خروجی برابر حداکثر ۱۲۸ در نظر گرفته شده‌اند) بنابراین نیازی به استفاده از همه این تعبیه نداشته و از بخش بسیار کوچک آن برای ساخت تعبیه مکانی کلمه استفاده می‌کنیم.

سوال سوم

قسمت الف

در مقاله Attention is all you need! پیشنهاد شده است که از بهینه‌ساز Adam با نرخ یادگیری که مقدار آن از طریق فرمول ۱ محاسبه می‌شود استفاده شود. به همین دلیل ما نیز از همین تابع خطا برای یادگیری استفاده می‌کنیم.

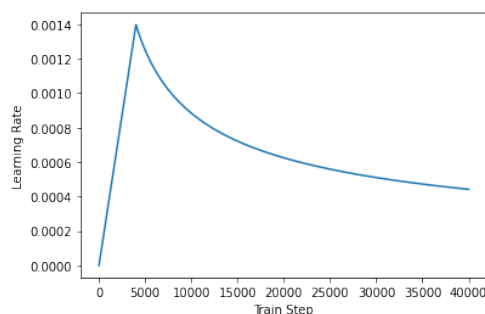
$$\eta = d_{model}^{-0.5} * \min(\text{step_num}^{-0.5}, \text{step_num} \cdot \text{warmup_step}^{-1.5}) \quad (1)$$

در شکل ۳ تغییرات نرخ یادگیری پیشنهاد شده در فرمول بالا در گام‌های یادگیری دیده می‌شود. در ابتدا نمودار صعودی بوده و مقدار نرخ یادگیری بیشتر و بیشتر می‌شود. با رسیدن به نقطه بیشینه که محل تلاقی خط $\frac{\text{step_num}}{\text{warmup_step}^{1.5}}$ و منحنی $\frac{1}{\text{step_num}^{-0.5}}$ است؛ شیب نمودار رو به پایین شده و مقدار نرخ یادگیری و در نتیجه آزادی عمل مدل رفته‌رفته کم‌کم‌تر می‌شود.

قسمت ب

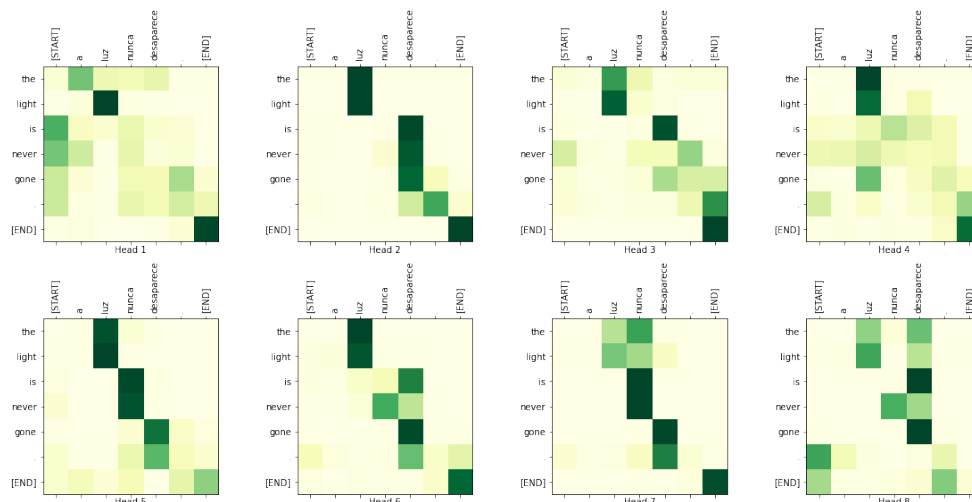
در شکل ۴ نحوه توجه مدل به ورودی‌ها برای تولید خروجی مشاهده می‌شود. در Head1 و Head4 مدل از یک دید کلی برای به دست آوردن تعبیه هر کلمه بر اساس متن بهره می‌گیرد.

در Head3 با توجه به آن که بیشتر توجه مدل روی دو کلمه light با کلمه متناظرش و کلمه is با ترجمه متناظرش است بنابراین هدف مدل از این توجه تطبیق دادن اسم و فعل است.



شکل ۳: تغییرات نرخ یادگیری بر اساس فرمول ۱

در Head2، Head5، Head6 و Head7 بیشتر توجه مدل روی گروه‌های اسمی و فعلی و به خصوص هسته این گروه‌هاست. مثلاً در Head2 مدل توجه زیادی به کلمه‌های *desparece* و ترجمه معادل آن یعنی *is never gone* کرده است. بعلاوه در همان Head2 مدل توجه زیادی به انتهای جمله داشته و به نظر می‌رسد که نمی‌خواسته طول ترجمه‌ای که ارائه می‌دهد از یک اندازه‌ای بیشتر باشد. در Head5 و Head6 مدل علاوه بر گروه‌های اصلی نظیر اسمی و فعلی به گروه قیدی *nunca* نیز توجه کرده و این مورد را در ترجمه دخیل کرده است.



شکل ۴: نقشه توجه مدل به ورودی‌ها

قسمت ج

در ادامه عملکرد مدل بر روی ۲۰ تا از جملات موجود در داده آزمون مشاهده می‌شود. با توجه به موارد پایین عملکرد مدل قابل است و اکثر ترجمه قابل قبول را ارائه داده است برای مثال مدل ۱۴ بر اساس ورودی که جمله خبری است خروجی را تولید کرده است؛ در حالی که ترجمه انسانی بر حسب مفهوم و سلیقه ترجمه را به شکل جمله

پرسشی انجام داده است. بعلاوه در بعضی موارد ترجمه خروجی دقیقاً جمله مرجع است. مثال ۵ یکی از این موارد است. دقت داشته باشید که از آن جا که این آزمون روی داده تست انجام شده است بنابراین یکسان بودن خروجی مدل و جمله مرجع بسیار حائز اهمیت است. بعضاً ایرادهایی نیز در عملکرد مدل مشاهده می‌شود. برای مثال در مثال شماره ۱۳ مدل نتوانسته است به خوبی قسمت how much negative energy it takes را ترجمه کند. یا مثلاً در مثال شماره ۱۰ مدل کلمه now را دوبار در خروجی تکرار کرده است.

Example 0

Input: : depois , podem fazer-se e testar-se previsões .
 Prediction : then you can make it and testify .
 Ground truth : then , predictions can be made and tested .

Example 1

Input: : forçou a parar múltiplos laboratórios que ofereciam testes brca .
 Prediction : it forced to stop multiple labs that they offer bruty test .
 Ground truth : it had forced multiple labs that were offering brca testing to stop .

Example 2

Input: : as formigas são um exemplo clássico ; as operárias trabalham para as rainhas e vice-versa .
 Prediction : the ants are a classic example ; operating in the queens and the queens will come .
 Ground truth : ants are a classic example ; workers work for queens and queens work for workers .

Example 3

Input: : uma em cada cem crianças no mundo nascem com uma doença cardíaca .
 Prediction : one in every 100 children in the world are born with a heart disease .
 Ground truth : one of every hundred children born worldwide has some kind of heart disease .

Example 4

Input: : neste momento da sua vida , ela está a sofrer de sida no seu expoente máximo e tinha pneumonia .
 Prediction : at this moment of his life , she ' s suffering from aids at his expopically , and she had pneumonia .
 Ground truth : at this point in her life , she ' s suffering with full-blown aids and had pneumonia .

Example 5

Input: : onde estão as redes económicas ?
 Prediction : where are economic networks ?
 Ground truth : where are economic networks ?

Example 6

Input: : (aplausos)
 Prediction : (applause)
 Ground truth : (applause)

Example 7

Input: : eu usei os contentores de transporte , e também os alunos ajudaram-nos a fazer toda a mobília dos edifícios , para torná-los confortáveis , dentro do orçamento do governo , mas também com a mesma
 Prediction : i used shipping containers , and they also helped us do all the way to nobilis of buildings , to make them comfortable , within the government , but also with the same area of the house , but suc
 Ground truth : i used the shipping container and also the students helped us to make all the building furniture to make them comfortable , within the budget of the government but also the area of the house is e

Example 8

Input: : e , no entanto , a ironia é que a única maneira de podermos fazer qualquer coisa nova é dar um passo nessa direção .
 Prediction : and yet the irony is that the only way we can do something new is to step one step towards that direction .
 Ground truth : and yet , the irony is , the only way we can ever do anything new is to step into that space .

Example 9

Input: : a luz nunca desaparece .
 Prediction : the light is never gone .
 Ground truth : the light never goes out .

Example 10

Input: : `` agora , `` '' tweets '' '' , quem está a `` '' tweetar '' '' ? ''
 Prediction : now , tweets now , who is tweeting ?
 Ground truth : now , tweets , who 's tweeting ?

Example 11

Input: : no egito , por exemplo , 91 % das mulheres que vivem hoje no egito foram mutiladas sexualmente dessa forma .
 Prediction : in egypt , for example , 91 percent of women who live today in egypt were mutilized by that way .
 Ground truth : in egypt , for instance , 91 percent of all the females that live in egypt today have been sexually mutilated in that way .

Example 12

Input: : por outro lado , os bebês de 15 meses ficavam a olhar para ela durante muito tempo caso ela agisse como se preferisse os brócolos , como se não percebessem a situação .
 Prediction : on the other hand , the babies were taking care for her for a long time , and they would fit in like to buy broccolicoilum , as if they did n ' t understand the situation .
 Ground truth : on the other hand , 15 month-olds would stare at her for a long time if she acted as if she liked the broccoli , like they could n't figure this out .

Example 13

Input: : naquele momento , percebi quanta energia negativa é precisa para conservar aquele ódio dentro de nós .
 Prediction : at that point , i realized how much power i do it is to constit that hate within us .
 Ground truth : in that instant , i realized how much negative energy it takes to hold that hatred inside of you .

Example 14

Input: : e a discussão é : o que é que isso interessa .
 Prediction : and the argument is : what ' s the point of it .
 Ground truth : and the discussion is , who cares ? right ?

Example 15

Input: : se escolhermos um lugar e formos cuidadosos , as coisas estarão sempre lá quando as procurarmos .
 Prediction : if you choose a place , and you go care , things will always be there when you look for them .
 Ground truth : if you designate a spot and you 're scrupulous about it , your things will always be there when you look for them .

Example 16

Input: : é um museu muito popular agora , e criei um grande monumento para o governo .
 Prediction : it ' s a very popular museum , and i made a great monumen for government .
 Ground truth : it 's a very popular museum now , and i created a big monument for the government .

Example 17

Input: : é completamente irrelevante .
 Prediction : it ' s completely irrelevant .
 Ground truth : it 's completely irrelevant .

Example 18

Input: : todos defenderam que a sua técnica era a melhor , mas nenhum deles tinha a certeza disso e admitiram-no .
 Prediction : all of them have advicted that their technique was the best , but no one of them was sure that it is admitted .
 Ground truth : `` they all argued that , `` '' my technique is the best , `` '' but none of them actually knew , and they admitted that . ''

Example 19

Input: : a partir daquele momento , comecei a pensar .
 Prediction : from that moment , i started thinking .
 Ground truth : at that moment , i started thinking .