

به نام خدا



دانشگاه صنعتی امیرکبیر
(پلی تکنیک تهران)

درس شبکه‌های عصبی
استاد صفابخش

تمرین هشتم

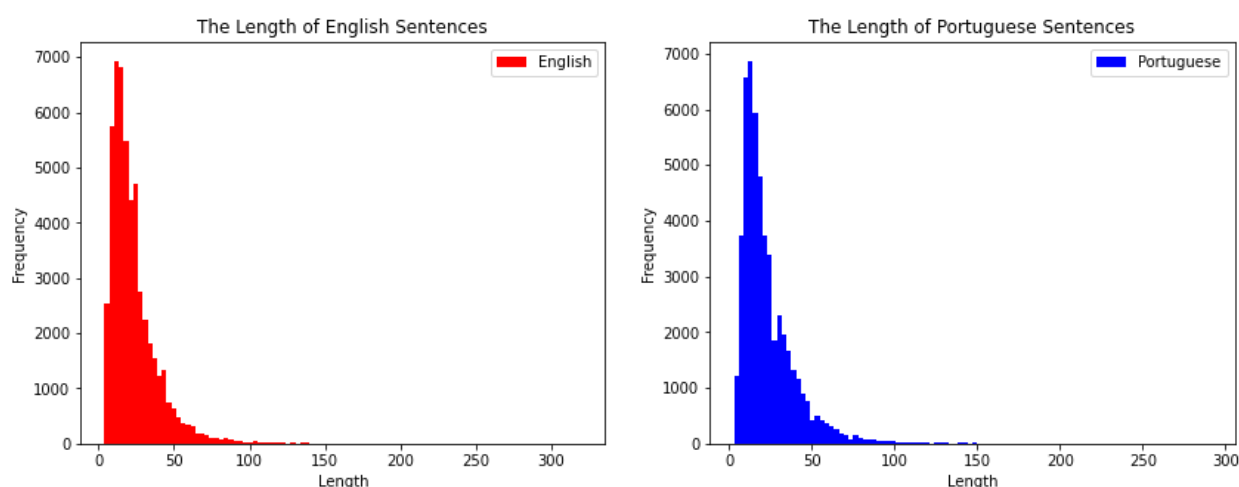
علیرضا مازوچی
۴۰۰۱۳۱۰۷۵

توجه

برای حل این تمرین از مطلب موجود در سایت تسورفلو^۱ استفاده شده است. از آنجایی که پیاده‌سازی موجود در این سایت کاملاً متناسب با این تمرین بوده است، شباهت زیادی بین پیاده‌سازی من با این پیاده‌سازی وجود دارد. عملاً منطق تمام مدل‌ها، پارامترهای مدل‌ها و بخشی از توابع پیچیده نظیر آموزش مدل عیناً و بدون تغییر مورد استفاده قرار گرفته است.

سوال ۱

برای پیش‌پردازش تصمیم گرفتم تا جملات بلند را حذف کنم. به این ترتیب امکان کاهش حافظه مصرفی وجود خواهد داشت. برای این کار باید طول حداکثری یک جمله را تعیین کنیم. این مقدار باید به گونه‌ای تعیین شود که اکثر جملات کوتاه‌تر از حد بالا باشند چراکه در غیر این صورت مجبور به حذف بخش مهمی از داده آموزشی خواهیم شد. در تصویر زیر مشخص است که برای زبان مبدا (پرتغالی) و مقصد (انگلیسی) هر جمله چه تعداد توکن دارد:



¹ <https://www.tensorflow.org/text/tutorials/transformer>

همانطو رکه از تصویر بر می‌آید برای هر دو زبان شرایط تقریباً مشابه است. بیشتر جملات هر دو زبان در حدود ۲۵ توکن دارند و به ندرت جمله‌ای را می‌توان یافت که بالای ۱۲۸ توکن داشته باشد. بنابراین حد بالای تعداد توکن در یک جمله را برابر با ۱۲۸ در نظر گرفته شده است.

سوال ۳ – بهینه‌ساز

مطابق با آموزش تسورفلو از بهینه‌ساز آدام با یک تابع برای تعیین مقدار نرخ یادگیری در هر لحظه استفاده کردم. این بهینه‌ساز در مقاله معروف Attention Is All You Need^۲ ارائه شده است. مطابق این مقاله پارامترهای آدام به این شکل کامل می‌شوند: $\beta_1 = 0.9$, $\beta_2 = 0.98$ و $\epsilon = 10^{-9}$. تابع انتخاب‌شده برای نرخ یادگیری به این شکل عمل می‌کند که ابتدا نرخ یادگیری تا یک یک تعداد گام که اصطلاحاً warm up steps نامیده می‌شود به صورت خطی زیاد می‌شود و سپس متناسب با معکوس مربعات ریشه شماره گام کاهش می‌یابد. تابع مقدار نرخ یادگیری به شکل دقیق عبارت است از:

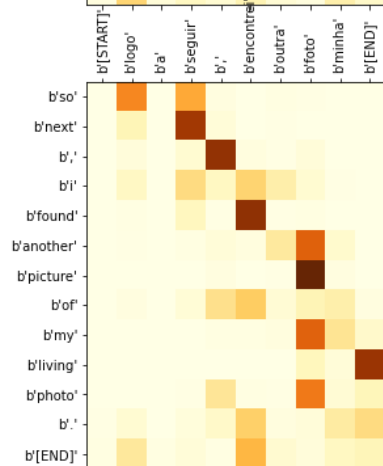
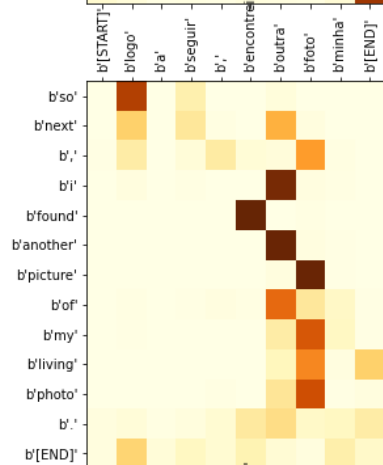
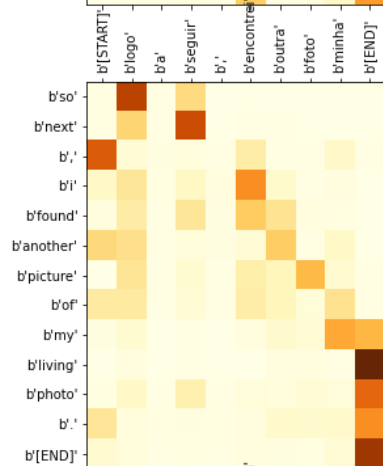
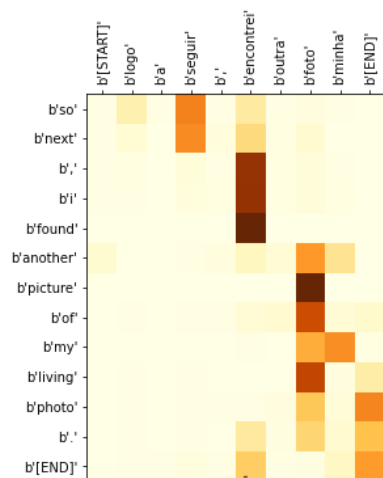
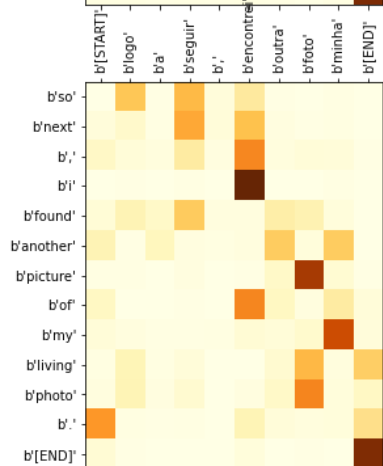
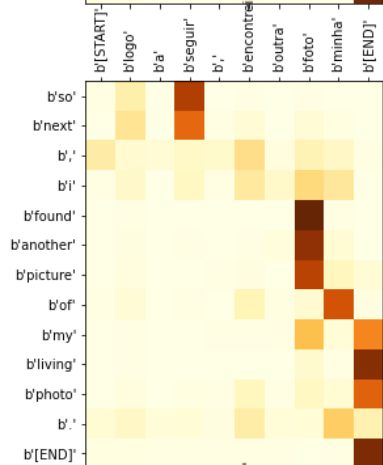
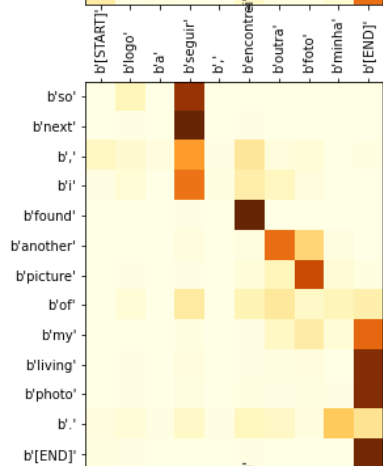
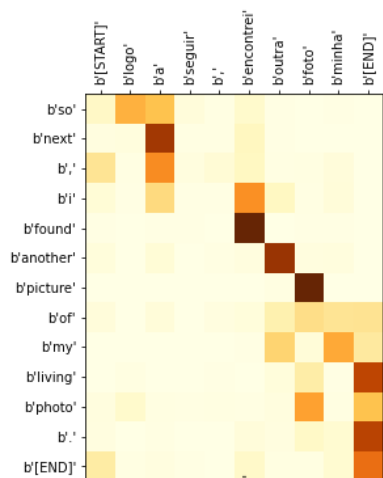
$$learning_rate(s) = d_model^{-0.5} * \min(s^{-0.5}, s * warm_up_steps^{-1.5})$$

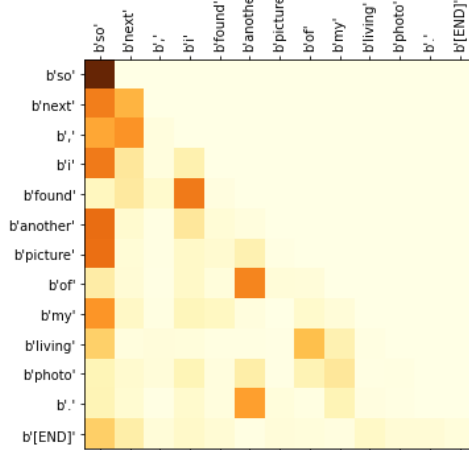
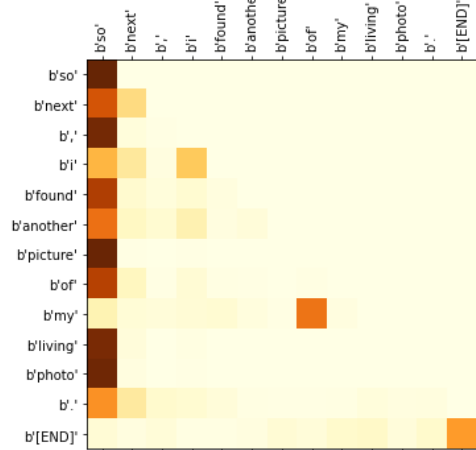
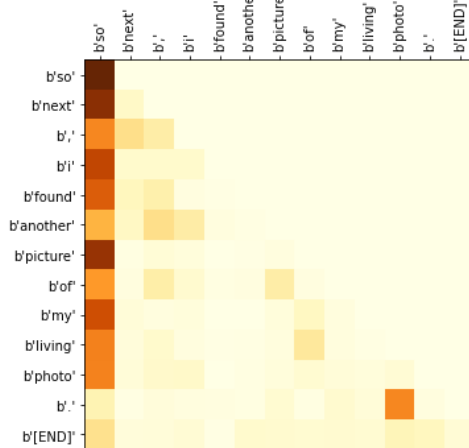
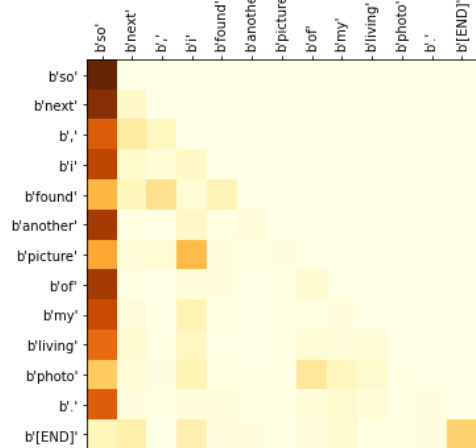
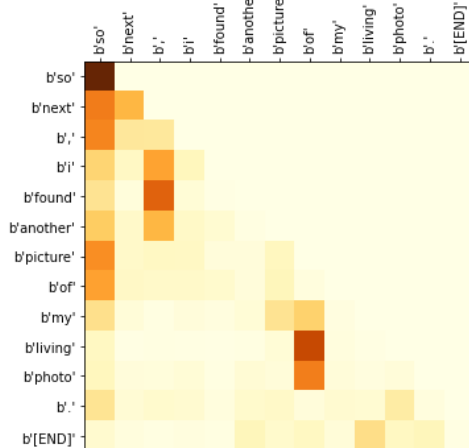
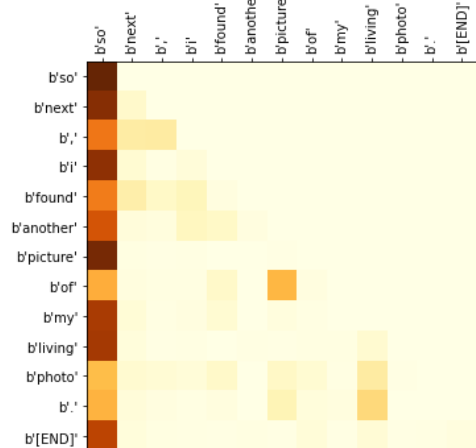
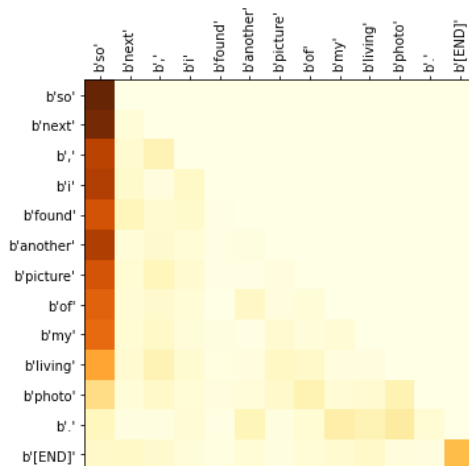
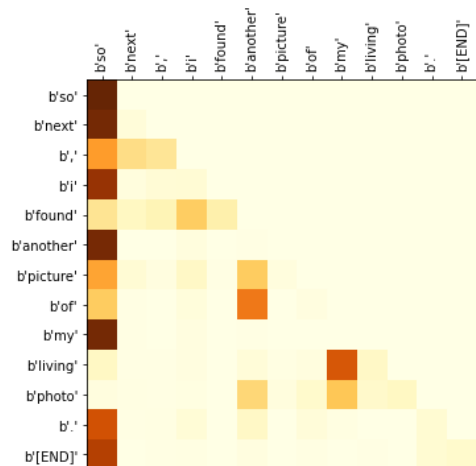
سوال ۳ – نقشه حرارتی توجه

برای تحلیل آسان‌تر تنها نقشه حرارتی توجه آخرین بلاک دیکودر در ادامه ترسیم می‌شود. هم توجه به خود دیکورد و هم توجه دیکودر بر اینکودر. جمله انتخابی هم عبارت است از «logo a seguir , encontrei outra foto minha» که ترجمه پیشنهادی مدل برای آن «so next , i found another picture of my living photo» است.

ابتدا توجه دیکودر بر اینکودر و سپس توجه به خود دیکودر آورده می‌شود:

² <https://arxiv.org/pdf/1706.03762.pdf>





برای تحلیل نتایج می‌توان دید هر کلمه مقصد در مجموع روی چه کلمه مبدا بیشتری تمرکز را داشته است. دو کلمه «so next» ترجمه «logo a seguir» هستند. در اکثر سرهای توجه این دو کلمه بیشترین توجه را به «logo» و «seguir» داشتند که منطقی است. سپس یک کاما در جمله انگلیسی دیده می‌شود که این کاما هم بیشترین توجه را به کلمات قبل، بعد و خود کاما در جمله پرتغالی داشته است. کلمه بعدی در جمله انگلیسی «i» است. فعل «encontrei» شامل ضمیر اول شخص است و طبیعی است که بیشترین توجه کلمه «i» بر روی همین کلمه باشد. کلمه بعدی فعل «found» است. ریشه و ضمیر این کلمه از روی فعل «encountrei» بدست می‌آید و باز مطابق انتظار بیشترین توجه را این فعل پرتغالی گرفته است. کلمه «another» ترجمه کلمه «outra» است که از روی نقشه حرارتی مشخص است. به طور مشابه کلمات «photo» و «picture» ترجمه «foto» هستند و شرایط مشابهی دارند. چند کلمه انتهایی بیشتر توجه بر روی خاتمه جمله داشتند و احتمالا سعی در این داشتند که متوجه شوند جمله به چه شکل باید تمام شود.

حال نتایج دومین نمودار را بررسی می‌کنیم. در این نمودار به طرز عجیبی روی اولین کلمه توجه زیادی وجود دارد. اگرچه برای چند کلمه اول چنین چیزی منطقی است ولی برای کلمات بعدی چندان منطقی به نظر نمی‌رسد. به هر حال چند تا از مواردی که توجه به غیر از کلمه اول بوده بررسی می‌گردد. در نمودار چپ اولین سطر می‌بینیم که کلمه «found» بر روی کلمه «i» توجه کرده است که این توجه احتمالا برای تعیین صیغه فعل بوده است. در همین نمودار «of» بر «another» توجه داشته است چراکه اگر حرف اضافه‌ای بخواهد بیاید باید با «another» متناسب باشد. نهایتا توجه زیادی از «living» به «my» می‌بینیم که کاملا طبیعی است چراکه باید ترکیب «my living» باهم سازگار باشد. در نمودار راست از سطر دوم مشاهده می‌شود که هر سه کلمه «my living photo» به «of» توجه داشتند که چون عبارت بعد از آن هستند موجه است.

سوال ۳ - نمونه خروجی

ردیف	جمله پرتغالی	جمله انگلیسی صحیح	جمله انگلیسی پیشنهادی
۱	depois , podem fazer-se e testar-se . previsões	then , predictions can be made and . tested	then you can make it and read . yourself predictions
۲	forçou a parar múltiplos laboratórios . que ofereciam testes brca	it had forced multiple labs that were . offering brca testing to stop	it forced to stop multiple labs which . managers to fight bruton
۳	as formigas são um exemplo clássico ; as operárias trabalham . para as rainhas e vice-versa	ants are a classic example ; workers work for queens and queens work for . workers	ants are an example : operating for . the queens and vice versa
۴	uma em cada cem crianças no mundo nascem com uma doença . cardíaca	one of every hundred children born worldwide has some kind of heart . disease	one of every hundred children in the . world are born with a heart disease
۵	neste momento da sua vida , ela está a sofrer de sida no seu expoente . máximo e tinha pneumonia	at this point in her life , she 's suffering with full-blown aids and had . pneumonia	at this moment of her life , she ' s suffering from aids in her exposure . and had pneumonia
۶	? onde estão as redes económicas	? where are economic networks	? where are the economic networks
۷	(aplausos)	(applause)	(applause)
۸	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 a área da . casa , mas muito mais confortável	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 exactly the same , . but much more comfortable	i used shipping , and the students also helped us to make the entire buildings of buildings , to make them comfortable , in the government budget of the government budget , but also with the same area at home , but they also have a lot more . comfortable with the house
۹	e , no entanto , a ironia é que a única maneira de podermos fazer qualquer coisa nova é dar um passo nessa . direção	and yet , the irony is , the only way we can ever do anything new is to . step into that space	and yet , the irony is that the only way we can do anything new and . take a step back to that direction
۱۰	. a luz nunca desaparece	. the light never goes out	. light has never gone
۱۱	agora , " " tweets " " , quem está " ? " " a " " tweetar	? now , tweets , who 's tweeting	now , tweets are , who is the ? twemplate
۱۲	no egito , por exemplo , 91 % das mulheres que vivem hoje no egito foram mutiladas sexualmente dessa . forma	in egypt , for instance , 91 percent of all the females that live in egypt today have been sexually mutilated in . that way	in egypt , for example , 91 percent of women who live today in egypt were sexually in the sexually in that way . that was sexually
۱۳	por outro lado , os bebés 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 .	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	on the other hand , the children of 15 months were going to look at her for a long time , and she took off as if she was in the brocreprofits , like not . understanding the situation
۱۴	naquele momento , percebi quanta energia negativa é precisa para conservar aquele ódio dentro de nós .	in that instant , i realized how much negative energy it takes to hold that . hatred inside of you	at that moment , i realized how much negative energy it needs to look at . that hatred in the place of us
۱۵	e a discussão é : o que é que isso . interessa	and the discussion is , who cares ? ? right	and the argument is : what that ' s . going on

if you pick a place , and you go and you ' re going to , the things will always be there when you look for . them	if you designate a spot and you 're scrupulous about it , your things will always be there when you look for . them	se escolhermos um lugar e formos cuidadosos , as coisas estarão . sempre lá quando as procurarmos	١٦
it ' s a very popular museum now , and i ' ve created a large monumen . for government	it 's a very popular museum now , and i created a big monument for the . government	é um museu muito popular agora , e criei um grande monumento para o . governo	١٧
. it ' s completely irreprepile	. it 's completely irrelevant	. é completamente irrelevante	١٨
they all defend that their technique was the best , but no one of them was certain of that they were . admitted and admitted it	they all argued that , `` " my `` technique is the best , " " but none of them actually knew , and they " . admitted that	todos defenderam que a sua técnica era a melhor , mas nenhum deles tinha a certeza disso e admitiram-no .	١٩
from that moment , i started thinking .	. at that moment , i started thinking	a partir daquele momento , comecei . a pensar	٢٠