



Prediction of missing words

HELLO!

Mohadeseh Jabbari

Fatemeh Tabadkani

Dr kamaledin Giasi

- Bert یک روش مبتنی بر ترانسفورمرها برای یادگیری بازنمایی زبان است.
- این یک bidirectional transformer از پیش آموزش دیده است که به لطف دو رویکرد منحصر به فرد، (mask language modeling (MLM و next sentence prediction (NSP از موفقیت بی نظیری در NLP برخوردار بوده است.
- ما میتوانیم از بازنمایی های زبانی که BERT آموخته است برای کارهای مانند طبقه بندی متن و... استفاده کنیم تا به نتایج پیشرفته ای در مورد مسئله خود دست یابیم.
- در بسیاری از موارد، ممکن است بتوانیم مدل bert را که از قبل آموزش داده شده است، خارج کنیم و آن را با توجه به مدل مورد نیاز در مسئله پیاده سازی کنیم.
- این ترنسفورمر که مورد توجه بسیاری از افراد قرار گرفت، برای شرکت گوگل ۷ هزار دلار هزینه دارد.

ما در واقع از Bert اینگونه استفاده میکنیم که یک جمله ناقص وارد میکنیم و از BERT میخواهیم که جمله را برای ما کامل کند.

به عنوان مثال در جمله زیر میخواهیم جای خالی را پر کنیم :

In Autumn the ----- fall from the trees.

در این جمله به احتمال زیاد جواب جای خالی را میدانید و این به این دلیل است که متن جمله را در نظر گرفته اید و مفهوم جمله را متوجه شده اید.

چیزهای زیادی از درختان می افتند به عنوان مثال بلوطها، شاخه ها، برگ ها. اما ما در پاییز شرایط دیگری داریم، که جستجوی ما را محدود می کند، محتمل ترین چیزی که در پاییز از درخت می افتد، برگ است.

Answer: leaves

ما به عنوان انسان ترکیبی از دانش عمومی و درک زبانی برای رسیدن به این نتیجه استفاده می کنیم. برای BERT این حدس از مطالعه زیاد و یادگیری الگوهای زبانی فوق العاده خوب حاصل می شود.

BERT ممکن است نداند پاییز، درختان و برگ ها چیست، اما می داند که با توجه به الگوهای زبانی و زمینه این کلمات، پاسخ به احتمال زیاد برگ ها است.

ما یک متن از کتاب Actual test آیلتس را انتخاب کرده ایم و فعل جملات هر پاراگراف را بدست آوردیم که به عنوان جای خالی استفاده کنیم و از bert کمک گرفته ایم تا جای خالی ها را برایمان پر کند.

جای خالی میتواند برای صفات، قیدها یا اسم ها نیز به کار برود.

جای خالی ها را در سه سطح بررسی کردیم: Basic, Intermediate, Advanced

Log In

Username:

Password:

Log In

If you already have no account, [register](#) instead.

Register

Register

Username:

Email:

Password:

Password confirmation:

Register

If you already have an account, [login](#) instead.

<http://jalammar.github.io/illustrated-transformer/>

<https://medium.com/analytics-vidhya/fine-tuning-bert-language-model-to-get-better-results-on-text-classification-3dac5e3c348e>

<https://towardsdatascience.com/masked-language-modelling-with-bert-7d49793e5d2c>

<https://github.com/Shivampanwar/Bert-text-classification>