

Evaluation Only. Created with Aspose.Words. Copyright 2003-2024 Aspose Pty Ltd.

-----INPUT-----:

Why Self-Attention?

-----OUTPUT-----:

SOURCE_0: .\lib\SOURCE_DOCUMENTS\Attention Is All You Need.pdf

CONTENT:

Self-attention, sometimes called intra-attention is an attention mechanism relating different positions of a single sequence in order to compute a representation of the sequence. Self-attention has been used successfully in a variety of tasks including reading comprehension, abstractive summarization, textual entailment and learning task-independent sentence representations [4, 27, 28, 22].

End-to-end memory networks are based on a recurrent attention mechanism instead of sequence-aligned recurrence and have been shown to perform well on simple-language question answering and language modeling tasks [34].

To the best of our knowledge, however, the Transformer is the first transduction model relying entirely on self-attention to compute representations of its input and output without using sequence-aligned RNNs or convolution. In the following sections, we will describe the Transformer, motivate self-attention and discuss its advantages over models such as [17, 18] and [9].

Similarity L2 Score: 0.26887554

SOURCE_1: .\lib\SOURCE_DOCUMENTS\Attention Is All You Need.pdf

CONTENT:

4 Why Self-Attention

In this section we compare various aspects of self-attention layers to the recurrent and convolutional layers commonly used for mapping one variable-length sequence of symbol representations (x_1, \dots, x_n) to another sequence of equal length (z_1, \dots, z_n) , with $x_i, z_i \in \mathbb{R}^d$, such as a hidden layer in a typical sequence transduction encoder or decoder. Motivating our use of self-attention we consider three desiderata.

One is the total computational complexity per layer. Another is the amount of computation that can be parallelized, as measured by the minimum number of sequential operations required.

Similarity L2 Score: 0.27875832

DATE: 2024-04-14 23:25:42.848171

-----INPUT-----:

What is the LoRA?

-----OUTPUT-----:

SOURCE_0: .\lib\SOURCE_DOCUMENTS\Easily Train a Specialized LLM.docx

CONTENT:

An ecosystem. LoRA is a practically useful tool that gives (almost) anyone the power to train a specialized LLM over their data. As a result, LoRA has been widely studied within the AI research community, leading to a variety of extensions, alternatives, and practical tools to go

Created with an evaluation copy of Aspose.Words. To discover the full versions of our APIs please visit: <https://products.aspose.com/words/>

along with it. One of the most notable extensions is QLoRa, which combines LoRA with model quantization to further reduce the memory overhead of LLM finetuning. However, this reduction in memory overhead comes at the cost of a slight decrease in training speed.

New to the newsletter?

Hi! I'm Cameron R. Wolfe, deep learning Ph.D. and Director of AI at Rebuy. This is the Deep (Learning) Focus newsletter, where I help readers understand AI research via overviews of relevant topics from the ground up. If you like the newsletter, please subscribe, share it, or follow me on Medium, X, and LinkedIn!

Bibliography

Similarity L2 Score: 0.27536142

SOURCE_1: .\lib\SOURCE_DOCUMENTS\Easily Train a Specialized LLM.docx

CONTENT:

Tied-LoRA [29]: leverages weight tying to further improve the parameter efficiency of LoRA.

GLoRA [30]: extends LoRA to adapt pretrained model weights and activations to each task in addition to an adapter for each layer.

Given that so many LoRA-inspired techniques exist, there are probably a few notable extensions that are missing from the list above. If you are aware of any other techniques that are worth including, let me know in the comments!

Takeaways

We should now have a working understanding of LoRA, the several variants of this technique that have been proposed, and how these ideas can be applied in practice. LoRA is arguably the most widely-used practical tool for creating specialized LLMs, as it democratizes the finetuning process by significantly reducing hardware requirements. Some important takeaways are outlined below.

Similarity L2 Score: 0.3139612

DATE: 2024-04-14 23:26:06.457274

-----INPUT-----:

Dokuz eylül üniversitesinde kaç adet fakülte var?

-----OUTPUT-----:

SOURCE_0: .lib\SOURCE_DOCUMENTS\DEU about.txt

CONTENT:

Dokuz Eylül Üniversitesi bugün; İzmir'in dört bir yanında; 10 Enstitü, 18 Fakülte, 2 Yüksekokul, 1 Konservatuvar, 6 Meslek Yüksekokulu, 2'si Uygulama ve Araştırma Hastanesi olmak üzere toplam 42 Uygulama ve Araştırma Merkezi ile bilimsel araştırma ve yükseköğretim görevini, kaliteli toplumsal hizmet anlayışı ile sürdürmektedir.

Dokuz Eylül Üniversitesi'nde toplam 72 bin 400'e yakın öğrenci eğitim görmektedir. Bunlar arasında, 114 ülkeden gelen 6 bin 300'den fazla uluslararası öğrenci bulunmaktadır. Üniversite bünyesinde 3 bine yakın akademik personel ve 4 bin 900'e yakın idari personel görev yapmaktadır

Similarity L2 Score: 0.2095406

SOURCE_1: .lib\SOURCE_DOCUMENTS\DEU about.txt

CONTENT:

Yine 2021 yılında Reha Midilli Foça Turizm Fakültesinin adı Turizm Fakültesi, Seferihisar Fevziye Hepkon Uygulamalı Bilimler Yüksekokulunun adı Uygulamalı Bilimler Yüksekokulu, Spor Bilimleri Fakültesinin adı Necat Hepkon Spor Bilimleri Fakültesi, Fizik Tedavi ve Rehabilitasyon Yüksekokulunun adı Fizik Tedavi ve Rehabilitasyon Fakültesi olarak değiştirilmiştir.

YÖK tarafından, 'Araştırma Kapasitesi', 'Araştırma Kalitesi' ve 'Etkileşim ve İş Birliği' başlıkları altındaki 32 gösterge kapsamında yapılan değerlendirmede gösterdiği yüksek performansla başarılı bir sınav vererek üniversitelerin 'şampiyonlar ligi'ne giren Dokuz Eylül Üniversitesi, 13.12.2021 tarihi itibarı ile sunduğu geniş araştırma olanakları, uluslararası ve ulusal projelerdeki başarısıyla "Araştırma Üniversitesi" unvanı almıştır.

Similarity L2 Score: 0.25700557

DATE: 2024-04-14 23:26:27.282183
