



**Cognitive
Science**

FACULTY OF PSYCHOLOGY UW
INSTITUTE OF PHILOSOPHY UW

Domain-specific text generation

Maria Bancerek

Peripatetic Conference X, Kalatówki, 21-24/10/2021

Text generation with language models is a crucial aspect of Artificial Intelligence. Fine-tuning a GTP-2 model enables us to design models that are sensitive to the generation context and exhibit high fluency in the task. Moreover, the fluency control measures enable us to reduce the fluency bias experienced by GTP-2 models. Natural language generation (NLG) tasks demand large amounts of highly descriptive texts, which requires a large amount of computational power, as well as high complexity to run. Specific datasets are often complex and the task of automatically generating end-to-end fluent utterances for a wide range of tasks is immense. Analyses on NLG data are useful for understanding the neural architecture of models and for evaluating the performance of encoders. Text evaluation for NLG systems has been extensively studied and evaluated. I discuss the shortcomings of NLG evaluation and highlight proposed solutions to these challenges.

Text generation with language models is a crucial aspect of Artificial Intelligence. Fine-tuning a GTP-2 model enables us to design models that are sensitive to the generation context and exhibit high fluency in the task. Moreover, the fluency control measures enable us to reduce the fluency bias experienced by GTP-2 models. Natural language generation (NLG) tasks demand large amounts of highly descriptive texts, which requires a large amount of computational power, as well as high complexity to run. Specific datasets are often complex and the task of automatically generating end-to-end fluent utterances for a wide range of tasks is immense. Analyses on NLG data are useful for understanding the neural architecture of models and for evaluating the performance of encoders. Text evaluation for NLG systems has been extensively studied and evaluated. I discuss the shortcomings of NLG evaluation and highlight proposed solutions to these challenges.

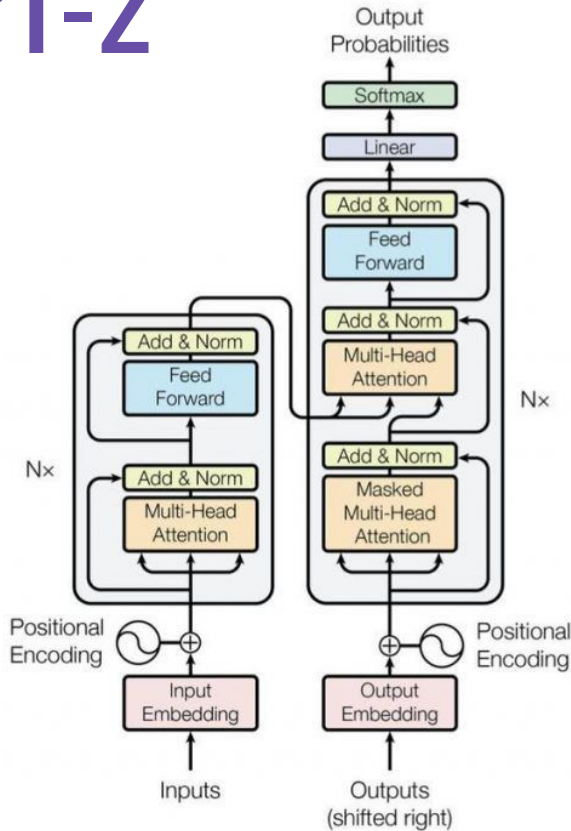


not written by me :) this is generated using a language model!

The model: GPT-2



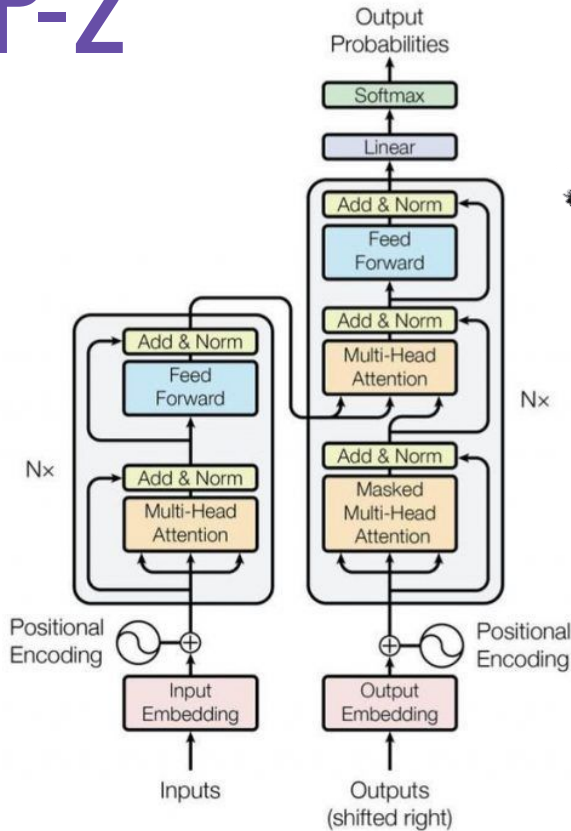
- Generative Pretrained Transformer (GPT-2) by OpenAI
- 1.5B parameters in original version
- trained on 40GB text data ~ 7 million A4 pages!
- smaller version (124M parameters)



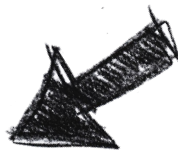
Finetuning GPT-2



- Generative Pretrained Transformer (GPT-2) by OpenAI
- 1.5B parameters in original version
- trained on 40GB text data ~ 7 million A4 pages!
- smaller version (124M parameters)



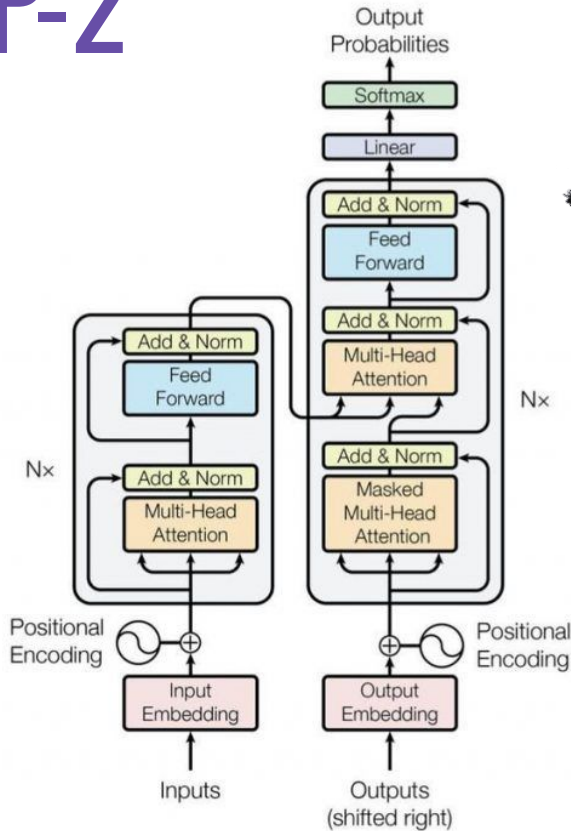
training on
specific dataset



Finetuning GPT-2



- Generative Pretrained Transformer (GPT-2) by OpenAI
- 1.5B parameters in original version
- trained on 40GB text data ~ 7 million A4 pages!
- smaller version (124M parameters)



training on
specific dataset

- 200 abstracts from arXiv (with a keyword, e.g. natural language generation)
- finetuning was adapted from: <https://github.com/minimaxir/gpt-2-simple>

Results

Exemplary abstract:

(Julian Zubek & Joanna Rączaszek-Leonardi)

Last 10 years have seen a proliferation of computational models of language emergence (Lazaridou & Baroni, 2020). Despite their variety (in concerns and architectures) they still seem to cling to several key assumptions about the communication process...

These include the belief that the semantic content of utterances matters, that the semantic content of propositional representations matters, and that the semantic content of utterances emerges from the relationships among the syntactic structure of the utterances and the semantic content of the speakers. These assumptions may be questioned, however, in light of the fact that there is very little empirical evidence to support the existence of implicit representations in language (Wu, 1986; Tamm, 1986; Wuhl, 1986; Gamm & Tamm, 1986).

We try to identify those among the assumptions which may hinder progress and point to possible ways to substitute them with less-limiting ones. We discuss three key issues. First, in the existing models environments tend to be extremely simplified, often reduced to a set of static stimuli presented to a passive agent. Agent's actions are limited to choosing the target stimulus from the set of distractors. There is no structure of agent's actions which could guide the structure of emerging communication.

Exemplary abstract:

(Julian Zubek & Joanna Rączaszek-Leonardi)

Last 10 years have seen a proliferation of computational models of language emergence (Lazaridou & Baroni, 2020). Despite their variety (in concerns and architectures) they still seem to cling to several key assumptions about the communication process...

These include the belief that the semantic content of utterances matters, that the semantic content of propositional representations matters, and that the semantic content of utterances emerges from the relationships among the syntactic structure of the utterances and the semantic content of the speakers. These assumptions may be questioned, however, in light of the fact that there is very little empirical evidence to support the existence of implicit representations in language (Wu, 1986; Tamm, 1986; Wuhl, 1986; Gamm & Tamm, 1986).

We try to identify those among the assumptions which may hinder progress and point to possible ways to substitute them with less-limiting ones. We discuss three key issues. First, in the existing models environments tend to be extremely simplified, often reduced to a set of static stimuli presented to a passive agent. Agent's actions are limited to choosing the target stimulus from the set of distractors. There is no structure of agent's actions which could guide the structure of emerging communication.

Typical artifacts

Last 10 years have seen a proliferation of computational models of language emergence (Lazaridou & Baroni, 2020). Despite their variety (in concerns and architectures) they still seem to cling to several key assumptions about the communication process...

First, there is a shared grammar. **Second**, there is a shared semantics. **Third**, there is a shared semantics for error correction (SDE). **Fourth**, there is a shared semantics for evaluating transformations. **Fifth**, there is a shared semantics for evaluating masking. **Sixth**, there is a shared semantics for evaluating lexical transformations. **Seventh**, there is a shared semantics for evaluating embedding. **Eighth**, there is a shared semantics for evaluating functions...

‘too regular’

Typical artifacts

Last 10 years have seen a proliferation of computational models of language emergence (Lazaridou & Baroni, 2020). Despite their variety (in concerns and architectures) they still seem to cling to several key assumptions about the communication process...

First, there is a shared grammar. **Second**, there is a shared semantics. **Third**, there is a shared semantics for error correction (SDE). **Fourth**, there is a shared semantics for evaluating transformations. **Fifth**, there is a shared semantics for evaluating masking. **Sixth**, there is a shared semantics for evaluating lexical transformations. **Seventh**, there is a shared semantics for evaluating embedding. **Eighth**, there is a shared semantics for evaluating functions...

These include the notion of context (i.e., the notion of what the author means), attention (i.e., the notion of what the author means) and the **emergence of new computational models**. In order to better understand **the emergence of new computational models**, we argue that models should focus on capturing **the emergence of new computational models** by considering **the emergence of new models** by...

‘too regular’

repetitions

A perfect peripatetic abstract?

The concept of entropy is derived from the concept of protein kinetics. In the real world our senses are complex and flexible and can perceive vast amounts of data. However, our perception of minimalism is a result of a cultural misunderstanding of how cognition is understood in humans. Prior to the emergence of digital communication, most people did not understand the concept of the entropic minimum necessary for achieving an end result. Large biological systems, such as humans, need a minimum of one genetic mutation to produce a child. (...) In complex systems such as societies or brains a significant degree of homology is usually made between the processes of emergence of language and the emergence of the other major properties.

Moral of the story: There is no unified perspective on this problem in philosophy. The position of a great many people is that the only truly effective form of communication is One-to-One Interference with Minds.

Peripatetic Conference X, Kalatówki, 21-24/10/2021

Thank you

Peripatetic Conference X, Kalatówki, 21-24/10/2021

Exemplary generated text:

Dance music is a popular activity for a group of high-functioning neural networks to dance. In this paper, we investigate the activity of neural networks trained on a domain-specific mixup of sample rates and settings to perform dance analysis. In this way, the neural networks are able to capture the dynamic range of the sound and determine the appropriate intensity level for the task...

A proposal to revise Britain's criminal code would put a high-speed internet cable over the top and put online a record of every internet pass through, the Mail on Sunday has revealed. The proposals would also make it easier to banish content overseas, as in a country like China, which has no plans to make a video service over the internet. Legal challenges have been filed at the Supreme Court...

So I thought, if I am going to make a living as a musician, I must be making a living as a musician'—she said. I can do it. I can learn any language I want at the school I want to go to, and I can sing English well. I have done it by doing what many good singers do in the afternoon: singing in the tones of cool credulity, rather than the cool, calculated, cowardly, cruel tongue...

Exemplary generated text:

Dance music is a popular activity for a group of high-functioning neural networks to dance. In this paper, we investigate the activity of neural networks trained on a domain-specific mixup of sample rates and settings to perform dance analysis. In this way, the neural networks are able to capture the dynamic range of the sound and determine the appropriate intensity level for the task...

NEWS

So I thought, if I am going to make a living as a musician, I must be making a living as a musician’—she said. I can do it. I can learn any language I want at the school I want to go to, and I can sing English well. I have done it by doing what many good singers do in the afternoon: singing in the tones of cool credulity, rather than the cool, calculated, cowardly, cruel tongue...

ABSTRACTS

A proposal to revise Britain’s criminal code would put a high-speed internet cable over the top and put online a record of every internet pass through, the Mail on Sunday has revealed. The proposals would also make it easier to banish content overseas, as in a country like China, which has no plans to make a video service over the internet. Legal challenges have been filed at the Supreme Court...

NOVELS

Finishing sentences

My eyes filled with tears...

abstracts:

My eyes filled with tears as I read these accounts of the channeling of deep neural networks with their non-stationary counterparts. As an experiment to test the theory, we also experiment with modeling neural network models (e.g., linear models) on image classification and Statistical Parameterization data. Our results show that, compared with the non-stationary models, the model with the more expensive DNN can reconstruct the label picture better. Additionally, we discovered that the model of DNNs that capture only the information...

news:

My eyes filled with tears, I was ready to go home, but I couldn't get home until I got my green card. So I went in the first class class I was in, and I saw that these are the only seats we had. And I said to myself, this is the only place we want to be. And if we could get a few more seats in the whole school, we would do that. And that's what motivated me to go into nursing.

novels:

My eyes filled with tears as she stood looking at the sky, and as she said: O lads! my brothers are dead! The world spun before her as though it had never ended, for the Spirit of the Lord had shown her still more the way of a survivor. She looked up and said to the sky: I do not know where you are but you'll never go back, for I have already done this to you: now you ask me again, I will give you my brother again.