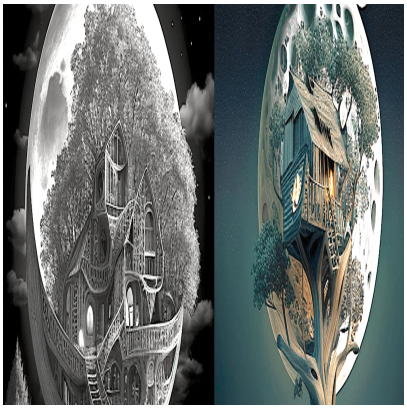

My tutor is an AI

Description

In constructing this post Iâ??ve been cross checking some of my explanations about neural networks with OpenAIâ??s chatGPT-3, which is a highly responsive chatbot available at the OpenAI.com website (free at present). Iâ??m most familiar with spreadsheets. So thatâ??s where Iâ??ll start.

You can think of a neural network as a matrix or grid (a kind of spreadsheet) in which every element (as the named rows and columns of a spreadsheet) has a numerical relationship with every other element. These numerical relationships are known as â??weights.â?• So, the cells in the spreadsheet are populated with numerical weight values. Each element also has its own value â?? a variable threshold value known as a â??bias.â?• Some of the elements in this spreadsheet function as inputs, some are outputs, and some are â??hidden.â?•

As we are talking about *networks*, itâ??s helpful to think of connected *nodes*: input nodes, output nodes and hidden nodes. Neural network designers also consider the input, output and hidden parts of the network as *layers*. There may be several hidden layers between the input and output layers. The nodes in a layer are heavily interconnected, effectively as input and output nodes connecting them to other layers.

Neural network algorithms perform calculations for each node. At its simplest, the process is to add the weights of all nodes connected to it (i.e. with non-zero values). If this summation is above the threshold value for that node then it â??firesâ?• meaning that it triggers the same process in nodes to which it is connected. To some extent this algorithmic process mirrors processes amongst neurons in organic neural systems (e.g. brains), though the processes in natural systems are massively parallel and interconnected.

Training

Algorithms that train a neural network make very small adjustments to weight and bias parameters across the whole network so that the network can reliably reproduce the same input and output patterns for each example in the training set. If the training is effective then each successive encounter with a new example preserves the ability of the network to reconstruct its earlier examples, i.e. there is little if

any degradation in the networkâ??s ability to reconstruct what it has â??learned.â?• Of even greater interest, over successive training events the network appears to behave as if it has inferred some general rules.

An early experimental neural network by Rumelhart and McLelland behaved as if it had learned the rules of grammar when trained with a large number of texts. The researchers showed that a network might improve its performance over successive training events, much as children might improve their grammatical performance over time, learning that the â??sâ?• at the end of a word makes it plural, but not to add an â??sâ?• to every plural word (overgeneralizing), and learning exceptions to rules. There are several methods and models for achieving effective machine learning with neural networks â?? for later discussion.

Hereâ??s how ChatGPT-3 explained the training process.

> The process of training a neural network involves both forward propagation, where the input data is passed through the network to produce an output, and backpropagation, where the error between the predicted output and the desired output is used to adjust the weights and biases of the network. This process is repeated for many input-output examples, and as the network â??learnsâ?• to produce the correct output for each example, it becomes better at generalizing to new, unseen examples.

Tokens

Whatever the training method, each input and output node of a neural network will correspond to an element in the domain the neural network is designed to deal with. Rather than â??elements,â?• neural network designers use the terms â??featuresâ?• or â??tokens.â?• Each input node corresponds to a *token* in the domain.

In a natural language processing system such as ChatGPT-3 the tokens are fragments of words. I asked ChatGPT-3 for examples of tokens that it uses. The chatbot tells me that its tokens are stored in a vast lookup table (known as a tokenizer) that includes â??50,257 unique tokens.â?•

> â??Here are some examples of tokens that you might find in the vocabulary of the GPT-3 tokenizer:
> Words: â??the,â?• â??and,â?• â??but,â?• â??that,â?• â??then,â?• â??when,â?• etc.
> Subwords: â??un,â?• â??der,â?• â??stan,â?• â??ding,â?• â??suc,â?• â??cess,â?• etc.
> Special characters: â??,â?• â??.â?• â??!â?• â???â?• â??â??â?• â??-â?• etc.â?•

The chatbot cautions:

> â??Keep in mind that the specific set of tokens in the GPT-3 vocabulary will depend on the specific version of the model and the data it was trained on. The tokens in the vocabulary represent the most frequently occurring units of text in the training data, so they will reflect the characteristics of that data.â?•

My interaction with ChatGPT-3 involves typing sentences into a chat field on the GPT-3 platform accessed through a web page at OpenAI.com. Before ChatGPT-3 can respond to my input with an output sequence, the GPT-3 tokenizer has to break my texts into tokens.

> â??The GPT-3 tokenizer is trained on a large dataset and uses this training data to learn how to split text into tokens in a way that is meaningful and useful for natural language processing tasks.â?•

By every measure, the performance of ChatGPT-3 is impressive. I show transcripts of previous interactions in earlier posts, e.g. Learn to talk to your AI. To give some indication of the scale of the platform, according to ChatGPT-3,

> â??the GPT-3 language model, [â?¦] was trained on a dataset of over 8 million web pages. GPT-3 is known for being one of the largest language models ever created, with 175 billion parameters.â?•

The parameters here are the weights and bias values in the layers of the neural network. They are the parameters that the neural network algorithm adjusts according to its machine learning model.

As I have already indicated, a neural network does not store input and output patterns but these emerge in the performance of the system. â??Learnedâ?• patterns are distributed across the network.

The crucial implication of the ability of these systems to generalise is that they are capable not only of reproducing what was in their original training set, but deliver as output new arrangements of tokens not in the original training set.

Input versus output

I asked ChatGPT-3 to provide an example of an input statement and the kind of output that it might generate.

> â??Input: â??I had a great time at the beach yesterday. The water was so clear and the sand was so soft.â?• Output: â??Wow, that sounds like a really nice day at the beach. I love going to the beach and relaxing in the sun. Did you do anything else besides just relax on the beach?â?•â?•

I asked what kind of source in the training set might have encouraged such an input-output pairing.

> â??It is likely that the input-output pair you provided was obtained from a conversation between two people. The input text (â??I had a great time at the beach yesterday. The water was so clear and the sand was so soft.â?•) could have been said by one person, and the output text (â??Wow, that sounds like a really nice day at the beach. I love going to the beach and relaxing in the sun. Did you do anything else besides just relax on the beach?â?•) could have been the response of the other person.â?•

(As the example is so well-formed and polite, my guess is that the training set included some text that includes this or similar examples to illustrate AI chat input and output.) The platform continued to inform me:

> â??The ChatGPT-3 model was trained on a large dataset of human conversations, so it is likely that this input-output pair was taken from one of those conversations. The conversations in the dataset likely covered a wide variety of topics and situations, so it is difficult to say exactly what kind of source this particular input-output pair came from.â?•

I then asked how ChatGPT-3 would deal with a statement somewhat removed from the training set, e.g. â??I had a great time on the moon yesterday. The air was so clear and the cheese was so soft.â?• The performance of the chatbot was such that it â??recalledâ?• what I had said earlier.

> â??It seems that you are inputting a statement about an experience on the moon. As I mentioned earlier, ChatGPT-3 is a language model that has been trained to generate human-like text based on the patterns and structures it learned from the data it was trained on. However, it is not able to generate text about experiences or events that are outside of the realm of what it was trained on.â?•

When I told it to treat this as an item of surreal poetry it replied with:

> â??The stars danced in the sky, spinning like glittering jewels on a cosmic carousel.â?•

I am here engaging in a meta-conversation (about a conversation) with a system deploying the methods I am trying to get it to explain. It explained that ChatGPT-3 is actually a version of the more general GPT-3 model that is not trained on specific input-output pairs:

> â??the GPT-3 model is designed to generate human-like text based on the patterns and structures it learned from the training data. It can be used to generate text that is similar to the text it was trained on.â?•

A useful tutor

Assuming ChatGPT-3â??s responses are accurate, it functioned for me as a useful tutor as well as an information source. On occasion I told it what I thought was a reasonable explanation of neural network functioning and it agreed or corrected me, politely. It seems just as helpful across other domains. See posts: Chatting with an AI about urban inexistence and The imitation game.

Bibliography

- Huang, Kelley. â??Alarmed by A.I. Chatbots, Universities Start Revamping How They Teach.â?• *New York Times*, January 16, 2023. Accessed 20 January 2023. https://archive.ph/2023.01.16-133026/https://www.nytimes.com/2023/01/16/technology/chatgpt-artificial-intelligence-

universities.html#selection-303.0-303.69
- Rumelhart, D.E. , and J.L. McClelland. â??On learning the past tense of English verbs.â?• In *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Volume 2: Psychology and Biological Models*, edited by J. L. McClelland, and D. E. Rumelhart, 216-271. Cambridge MA: MIT Press, 1987.
- Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. â??Attention Is All You Need.â?• In 31st Conference on Neural Information Processing Systems, 1-15. Long Beach, CA, USA. 2017.

Notes

- Transcript of my conversation with ChatGPT-3: PDF.
- Featured image is by Midjourney, prompted with â??Escher-style tree house on the moon.â?•

Category

1. Artificial Intelligence

Tags

1. chatgpt
2. neural networks

Date Created
January 21, 2023
Author
rcoyne99