GPT3 Info

* Generative pre-training
* “I am open to the idea that a worm with 302 neurons is conscious, so I am open to the idea that GPT-3 with 175 billion parameters is conscious too.” — David Chalmers
* Transformer model
  + Sequence to sequence deep learning models that can produce a sequence of text given an input speech
  + Designed for text generation tasks – answering questions, text summarisations and machine translation
  + Created as a neural network
* Extremely powerful without understanding anything it produces
* Multiple units called attention blocks to learn what part of a text sequence is important and what to focus on
  + Gpt3 has 96 attention blocks which each contain 96 attention heads
* Third gen model by openai released in June 2020 and tested for $14mil
* 17 times larger than gpt2 (175 bil parameters) and 10 times more than microsoft turing nlg model
  + Parameters are network calculations that apply particular weights to different aspects of data – each data receives its value and data perspective
    - Due to this, the language is capable of meta-learning – meaning gpt3 can do tasks without any training even from a single example
* Gpt3 trained using:
  + Common crawl
  + Webtext2
  + Books1
  + Books2
  + Wikipedia Corpus
* Entire startups created by using gpt3
* Currently not open-source – currently a private beta where people can sign up on a waitlist
  + Also offered as an API accessible through the cloud and some companies have developed some intriguing applications that use the generation of text to enhance all kinds of programs
* Recent experiments prove that it can simplify the work of developers by producing custom code
  + Python, CSS, JSX
  + Learning coding languages may become unnecessary as there will be one coding language for all cases
  + Could reduce skills required to be a programmer
  + Could also grow productivity and erase the need for low-skilled engineers
* Unlike BERT, gpt3 does not need examples and days to train and learn

**Issues with GPT-3**

* Lacks long term memory
* Lack of interpretability due to how big gpt3 is
* Limited input size – transformers fixed max input size. Meaning prompts can’t be longer than a few sentences
* Slow inference time because gpt3 is so big
* Suffers from bias – e.g. containing anti-muslim bias
* Can lack profound understanding of context
* Lacks human understanding and common sense

Diagram

Description automatically generatedGraphical user interface, application

Description automatically generated