

PLAY AROUND



- **BabyLM Challenge**
(<https://babylm.github.io/>)
- **Andrej Karpathy's build GPT from scratch**
(<https://www.youtube.com/watch?v=kCc8FmEb1nY>)
- **Inception interpretability**
(<https://distill.pub/2020/circuits/zoom-in/>)
- **GPT in 60 lines of NumPy**
(<https://jaykmody.com/blog/gpt-from-scratch/>)
- **PromptHero**
(<https://prompthero.com/>)

```

937 → 'first'
767 → 'man'
373 → 'on'
278 → 'the'
18786 → 'moon'
471 → 'was'
29871 → ' '

sampling parameters: temp = 0.800000, top_k = 40, top_p = 0.950000, repeat_last_n = 64, repeat_penalty = 1.300000

The first man on the moon was 20 years old and looked^[]

top - 18:16:11 up 147 days, 9:22, 5 users, load average: 9.45, 8.06, 5.11
tasks: 240 total, 2 running, 238 sleeping, 0 stopped, 0 zombie
%Cpu0 : 78.6 us, 7.8 sy, 0.0 ni, 0.0 id, 2.9 wa, 0.0 hi, 10.7 si, 0.0 st
%Cpu1 : 79.2 us, 13.2 sy, 0.0 ni, 0.0 id, 7.5 wa, 0.0 hi, 0.0 si, 0.0 st
%Cpu2 : 75.2 us, 16.2 sy, 0.0 ni, 0.0 id, 8.6 wa, 0.0 hi, 0.0 si, 0.0 st
%Cpu3 : 78.3 us, 13.2 sy, 0.0 ni, 0.0 id, 8.5 wa, 0.0 hi, 0.0 si, 0.0 st
MiB Mem : 3792.3 total, 83.5 free, 3621.3 used, 87.5 buff/cache
MiB Swap: 65536.0 total, 60299.7 free, 5236.2 used, 46.4 avail Mem

  PID USER      PR  NI   VIRT    RES    SHR S  %CPU  %MEM    TIME+   COMMAND
2705518 ubuntu    20   0 5231264    3.3g  1904 R 352.9   88.4   27:37.52  main
    102 root      20   0         0        0      0 S  12.5    0.0   28:11.15  kswapd0

Features                : fp asimd evtstrm crc32 cpuid
CPU implementer         : 0x41
CPU architecture        : 8
CPU variant              : 0x0
CPU part                 : 0xd08
CPU revision             : 3

Hardware                 : BCM2835
Revision                 : c03111
Serial                   : 10000000d62b612e
Model                    : Raspberry Pi 4 Model B Rev 1.1
ubuntu@rpi:~$
[0] 0:mc* 1:mc-

```

REFERENCES

- Stephen Wolfram's article on ChatGPT
(<https://writings.stephenwolfram.com/2023/02/what-is-chatgpt-doing-and-why-does-it-work/>)
- Attention is all you need
(<https://www.youtube.com/watch?v=iDulhoQ2pro>)
- Intro to Q learning
(<https://blog.floydhub.com/an-introduction-to-q-learning-reinforcement-learning/>)
- The illustrated transformer
(<http://jalammar.github.io/illustrated-transformer/>)
- Hugging Face NLP course
(<https://huggingface.co/course/chapter1/1>)
- How Transformers work in deep learning and NLP: an intuitive introduction
(<https://theaisummer.com/transformer/>)