Language Models are Few-Shot Learners
Ben Mann

2020.07.24
Zero shot

1. Translate English to French:

2. cheese => ...........................................

(task description) => (prompt)
Translate English to French:

1. sea otter => loutre de mer
2. peppermint => menthe poivrée
3. plush giraffe => girafe peluche
4. cheese => ........................................

Few shot
Finetuning

- sea otter => loutre de mer (example #1)
  - gradient update
    - peppermint => menthe poivrée (example #2)
      - gradient update
        - ...
          - plush giraffe => girafe peluche (example #N)
            - gradient update
              - cheese => ................ (prompt)
<table>
<thead>
<tr>
<th>Model Name</th>
<th>$n_{\text{params}}$</th>
<th>$n_{\text{layers}}$</th>
<th>$d_{\text{model}}$</th>
<th>$n_{\text{heads}}$</th>
<th>$d_{\text{head}}$</th>
<th>Batch Size</th>
<th>Learning Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPT-3 Small</td>
<td>125M</td>
<td>12</td>
<td>768</td>
<td>12</td>
<td>64</td>
<td>0.5M</td>
<td>$6.0 \times 10^{-4}$</td>
</tr>
<tr>
<td>GPT-3 Medium</td>
<td>350M</td>
<td>24</td>
<td>1024</td>
<td>16</td>
<td>64</td>
<td>0.5M</td>
<td>$3.0 \times 10^{-4}$</td>
</tr>
<tr>
<td>GPT-3 Large</td>
<td>760M</td>
<td>24</td>
<td>1536</td>
<td>16</td>
<td>96</td>
<td>0.5M</td>
<td>$2.5 \times 10^{-4}$</td>
</tr>
<tr>
<td>GPT-3 XL</td>
<td>1.3B</td>
<td>24</td>
<td>2048</td>
<td>24</td>
<td>128</td>
<td>1M</td>
<td>$2.0 \times 10^{-4}$</td>
</tr>
<tr>
<td>GPT-3 2.7B</td>
<td>2.7B</td>
<td>32</td>
<td>2560</td>
<td>32</td>
<td>80</td>
<td>1M</td>
<td>$1.6 \times 10^{-4}$</td>
</tr>
<tr>
<td>GPT-3 6.7B</td>
<td>6.7B</td>
<td>32</td>
<td>4096</td>
<td>32</td>
<td>128</td>
<td>2M</td>
<td>$1.2 \times 10^{-4}$</td>
</tr>
<tr>
<td>GPT-3 13B</td>
<td>13.0B</td>
<td>40</td>
<td>5140</td>
<td>40</td>
<td>128</td>
<td>2M</td>
<td>$1.0 \times 10^{-4}$</td>
</tr>
<tr>
<td>GPT-3 175B or “GPT-3”</td>
<td>175.0B</td>
<td>96</td>
<td>12288</td>
<td>96</td>
<td>128</td>
<td>3.2M</td>
<td>$0.6 \times 10^{-4}$</td>
</tr>
</tbody>
</table>
## Datasets

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Quantity (tokens)</th>
<th>Weight in training mix</th>
<th>Epochs elapsed when training for 300B tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>Common Crawl (filtered)</td>
<td>410 billion</td>
<td>60%</td>
<td>0.44</td>
</tr>
<tr>
<td>WebText2</td>
<td>19 billion</td>
<td>22%</td>
<td>2.9</td>
</tr>
<tr>
<td>Books1</td>
<td>12 billion</td>
<td>8%</td>
<td>1.9</td>
</tr>
<tr>
<td>Books2</td>
<td>55 billion</td>
<td>8%</td>
<td>0.43</td>
</tr>
<tr>
<td>Wikipedia</td>
<td>3 billion</td>
<td>3%</td>
<td>3.4</td>
</tr>
</tbody>
</table>
$L = 2.57 \cdot C^{-0.048}$
Human ability to detect model generated news articles

- Control (86%)
- Random chance (50%)

Accuracy (%) vs. Number of parameters (log scale)
Measuring and Preventing Memorization Of Benchmarks

- Train/test overlap
- Removal bug
- Clean vs dirty evals
- What does it mean to have a clean training set?
Limitations

- Perf/FLOP still low
- Poor multilingual performance
- Query entailment doesn't perform
- Didn't publish fine-tuning results
- "common sense"
- Reliability and bias
- Dataset quality
  - Extraction
  - Filtering
- Cost & deployability
Poor English input:  I eated the purple berries.
Good English output:  I ate the purple berries.

Poor English input:  Thank you for picking me as your designer.  I’d appreciate it.
Good English output:  Thank you for choosing me as your designer.  I appreciate it.

Poor English input:  The mentioned changes have done.  or I did the alteration that you requested.  or I changed things you wanted and did the modifications.
Good English output:  The requested changes have been made.  or I made the alteration that you requested.  or I changed things you wanted and made the modifications.

Poor English input:  I’d be more than happy to work with you in another project.
Good English output:  I’d be more than happy to work with you on another project.

Poor English input:  Please provide me with a short brief of the design you’re looking for and that’d be nice if you could share some examples or project you did before.
Good English output:  Please provide me with a brief description of the design you’re looking for and that would be nice if you could share some examples or projects you have done before.

Poor English input:  The patient was died.
Good English output:  The patient died.
Thank You

Visit openai.com for more information.

FOLLOW @OPENAI ON TWITTER