

Sabiá-7B Instruction Following



Antonio Isaac, Douglas Veras e Jonas Gabriel

Sumário

01

Apresentação
do problema

02

Objetivos

03

Dados utilizados e
pré-processamento
dos dados

04

Rede neural

05

Resultados



✦ Apresentação do problema ✦

- Sabia-7B
 - LLM desenvolvida pela empresa MaritacaAI
 - Baseado no LLaMA-7B
- Modelo pré-treinado disponível no Hugging Face



★ Apresentação do problema ★

- Sabia-7B possui pouca capacidade para seguir instruções humanas

```
[8] prompt = ""Coloque 3 nomes de brasileiros""
```

```
▶ input_ids = tokenizer(prompt, return_tensors="pt")

output = model.generate(
    input_ids["input_ids"].to("cuda"),
    max_length=1024,
    eos_token_id=tokenizer.encode("\n")) # Stop generation when a "\n" token is detected

# The output contains the input tokens, so we have to skip them.
output = output[0][len(input_ids["input_ids"][0]):]
```

```
[10] print(tokenizer.decode(output, skip_special_tokens=True))
```

```
famosos em ordem alfabética
```

★ Apresentação do problema ★

- Funciona para casos específicos e bem detalhados

```
prompt = """Classifique a resenha de filme como "positiva" ou "negativa".
```

```
Resenha: Gostei muito do filme, é o melhor do ano!
```

```
Classe: positiva
```

```
Resenha: O filme deixa muito a desejar.
```

```
Classe: negativa
```

```
Resenha: Apesar de longo, valeu o ingresso.
```

```
Classe: """
```

```
print(tokenizer.decode(output, skip_special_tokens=True))
```

```
positiva
```

★ Objetivos ★

- Avaliar a influência do refinamento do modelo base do Sabia-7B com um dataset de instruções em português
 - Até que ponto os dados incluídos melhoram a capacidade do modelo de seguir instruções



★ Dados utilizados e pré-processamento dos dados ★

- Dataset Canarim-Instruct-PTBR
 - Mais de 300.000 instruções em português.
 - Tarefas simples a complexas
 - Múltiplas fontes
- Adição de Tokens Especiais
 - <PAD> e </s>
- Tokenização
 - LlamaTokenizer (BPE)

★ Dados utilizados e pré-processamento dos dados ★

- Formatação dos dados para um formato de prompt com identificadores de:
 - Introdução de uma instrução;
 - A introdução em si;
 - Contexto, se houver;
 - Resposta adequada

★ Dados utilizados e pré-processamento dos dados.★

```
print(prompted_dataset['train'][1234]['prompt'])
```

Abaixo está uma instrução que descreve uma tarefa, emparelhada com uma entrada que fornece mais contexto. Escreva uma resposta que conclua adequadamente a solicitação.

Instrução:

Nomeie as capitais dos três países seguintes

Entrada:

Índia, Canadá, Egito

Resposta:

★ Dados utilizados e pré-processamento dos dados ★

```
print(prompted_dataset['train'][0]['prompt'])
```

Abaixo está uma instrução que descreve uma tarefa. Escreva uma resposta que conclua adequadamente a solicitação.

Instrução:

Dê três dicas para se manter saudável.

Resposta:

✧ Rede Neural. ✧

- Sabia-7B base
 - Mesma arquitetura do LLaMA-1-7B (Transformers)
- Parameter Efficient Fine Tuning (PEFT)
- LoRA e QLoRA

★ PEFT. ★

- Não faz um fine tuning completo dos pesos do modelo base
- “Congelamento” de alguns pesos
 - Treina os demais
- Menos recurso computacional
- Evita de perder informações previamente aprendidas



★ PEFT. ★

```
[ ] print(trainable_parameters(peft_model))
```

```
trainable model parameters: 12582912  
all model parameters: 3512995840  
percentage of trainable model parameters:  
    0.36%
```

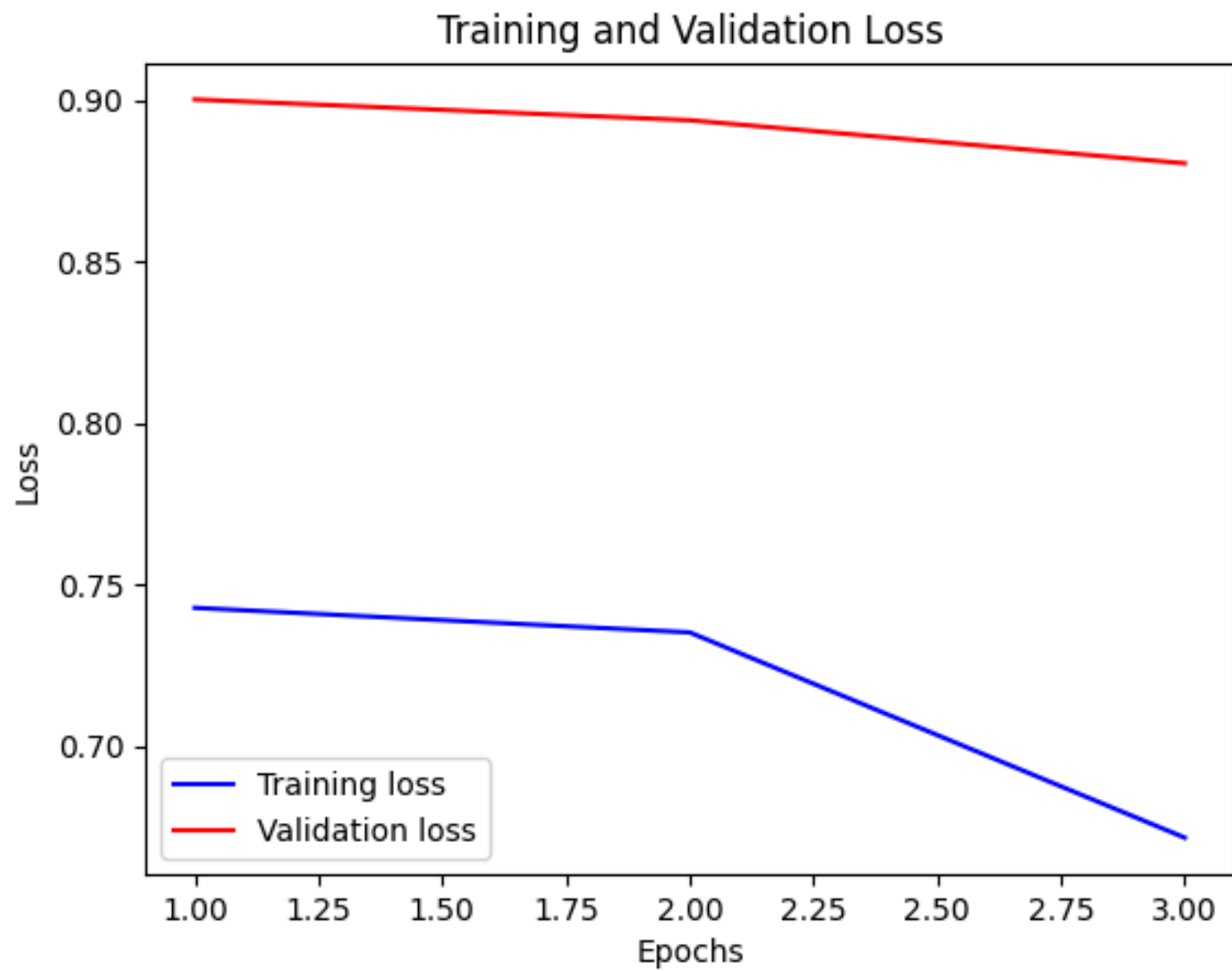
✦ LoRA e QLoRA ✦

- Local Rank Adaptation (LoRA)
 - Decompor a matriz de pesos originais em matrizes menores
 - Otimizar ainda mais os cálculos de treino
- QLoRA (Quantized Local Rank Adaptation)
 - Mudar a precisão dos pesos para 4 bytes

★ Treinamento. ★

- 27.000 textos de entrada para treino
- 1.500 para validação
- 3 Épocas
- Google Colab Pro
 - A100 GPU com 40GB de VRAM

★ Treinamento. ★



★ Resultados. ★

- O resultado obtido com o modelo foi razoável, porém com um problema
 - Saída da resposta sendo repetida a partir de um certo momento
- Não houve reconhecimento adequado do fim da sentença com o treinamento feito no modelo
 - Treinamento mais longo provavelmente resolveria esse problema.

★ Resultados. ★

- Falta de tempo e recursos para calcular métricas de avaliação para LLMs
 - Rouge
 - BLEU
 - Perplexity

Resultados.

[illegible]


The image is a screenshot of a web interface for a language model named Sabia7B Instruct. At the top, the name "Sabia7B Instruct" is displayed in a large, bold, white font, followed by a small red cardinal bird icon. Below this, there is a list of prompts and their corresponding outputs, each preceded by a yellow icon representing a prompt type (a shopping bag for general prompts and a face with a speech bubble for specific tasks). The prompts are: "Como posso te ajudar?", "Me de 3 comidas", "me de 3 nomes", and "John - Mary - Jane - Bill - Bob - Sally - Sue - Jack - Jill - Jim - Jane". The outputs are: "Pizza - Hambúrguer - Sushi - Frango - Carne - Peixe - Vegetariano - Salada - Omelete - Café da manhã - Lanche - Almoço - Jantar - Sobremesa - Bebida - Suco - Água - Cerveja - Vinho - Café - Chá - Chocolate - Leite - Iogurte - Manteiga - Queijo - Açúcar" for the second prompt, and "John - Mary - Jane - Bill - Bob - Sally - Sue - Jack - Jill - Jim - Jane" for the fourth prompt. At the bottom, there is a dark gray input field with the placeholder text "Diga algo" and a white arrow pointing to the right.


Resultados

Chá - Chocolate - Leite - Iogurte - Manteiga - Queijo - Açúcar

 me de 3 nomes

[illegible]

 o que é inteligencia artificial


 Inteligência artificial é a capacidade de um computador para pensar, aprender e tomar decisões como um ser humano. É uma área de pesquisa que estuda a criação de sistemas que podem imitar a inteligência humana. A IA é usada em muitos campos, como jogos, robótica, automação, reconhecimento de fala, reconhecimento de imagem e processamento de linguagem natural. A IA também pode ser usada para melhorar a eficiência de processos e tomar decisões mais precisas. A IA é uma área de pesquisa que está crescendo rapidamente, e muitos pesquisadores estão trabalhando para desenvolver novas tecnologias e aplicativos. A IA também está sendo usada em muitas indústrias, como finanças, saúde, educação e transporte. A IA é uma área de pesquisa que está crescendo rapidamente, e muitos pesquisadores estão trabalhando para desenvolver novas tecnologias e aplicativos. A IA também está sendo usada em muitas indústrias, como finanças, saúde, educação e transporte. A IA é uma área de pesquisa que está crescendo rapidamente, e muitos pesquisadores estão trabalhando para desenvolver novas tecnologias e aplicativos. A IA também está sendo usada em muitas indústrias, como finanças, saúde, educação e transporte. A IA é uma área de pesquisa que está crescendo rapidamente, e muitos pesquisadores estão trabalhando para desenvolver novas tecnologias e aplicativos.

A

 Como posso te ajudar?


 Me de 3 comidas

- Pizza - Hambúrguer - Sushi - Frango - Carne - Peixe - Vegetariano - Salada - Omelete - Café da manhã - Lanche - Almoço - Jantar - Sobremesa - Bebida - Suco - Água - Cerveja - Vinho - Café - Chá - Chocolate - Leite - Iogurte - Manteiga - Queijo - Açúcar

 me de 3 nomes

[illegible]

 o que é inteligencia artificial

 Inteligência artificial é a capacidade de um computador para pensar, aprender e tomar decisões como um ser humano. É uma área de pesquisa que estuda a criação de sistemas que podem imitar a inteligência humana. A IA é usada em muitos campos, como jogos, robótica, automação, reconhecimento de fala, reconhecimento de imagem e processamento de linguagem natural. A IA também pode ser usada para melhorar a eficiência de processos e tomar decisões mais precisas. A IA é uma área de pesquisa que está crescendo rapidamente, e muitos pesquisadores estão trabalhando para desenvolver novas tecnologias e aplicativos. A IA também está sendo usada em muitas indústrias, como finanças, saúde, educação e transporte.

7

Resultados.



o que é inteligência artificial



Inteligência artificial é a capacidade de um computador para pensar, aprender e tomar decisões como um ser humano. É uma área de pesquisa que estuda a criação de sistemas que podem imitar a inteligência humana. A IA é usada em muitos campos, como jogos, robótica, automação, reconhecimento de fala, reconhecimento de imagem e processamento de linguagem natural. A IA também pode ser usada para melhorar a eficiência de processos e tomar decisões mais precisas. A IA é uma área de pesquisa que está crescendo rapidamente, e muitos pesquisadores estão trabalhando para desenvolver novas tecnologias e aplicativos. A IA também está sendo usada em muitas indústrias, como finanças, saúde, educação e transporte. A IA é uma área de pesquisa que está crescendo rapidamente, e muitos pesquisadores estão trabalhando para desenvolver novas tecnologias e aplicativos. A IA também está sendo usada em muitas indústrias, como finanças, saúde, educação e transporte. A IA é uma área de pesquisa que está crescendo rapidamente, e muitos pesquisadores estão trabalhando para desenvolver novas tecnologias e aplicativos. A IA também está sendo usada em muitas indústrias, como finanças, saúde, educação e transporte. A IA é uma área de pesquisa que está crescendo rapidamente, e muitos pesquisadores estão trabalhando para desenvolver novas tecnologias e aplicativos.



O que é a vida

[illegible]

o que é inteligencia artificial



A inteligência artificial é a capacidade de um computador para pensar, aprender e tomar decisões como um ser humano. É uma área de pesquisa que estuda a criação de sistemas que podem imitar a inteligência humana. A IA é usada em muitos campos, como jogos, robótica, automação, reconhecimento de fala, reconhecimento de imagem e processamento de linguagem natural. A IA também pode ser usada para melhorar a eficiência de processos e tomar decisões mais precisas. A IA é uma área de pesquisa que está crescendo rapidamente, e muitos pesquisadores estão trabalhando para desenvolver novas tecnologias e aplicativos. A IA também está sendo usada em muitas indústrias, como finanças, saúde, educação e transporte. A IA é uma área de pesquisa que está crescendo rapidamente, e muitos pesquisadores estão trabalhando para desenvolver novas tecnologias e aplicativos. A IA também está sendo usada em muitas indústrias, como finanças, saúde, educação e transporte. A IA é uma área de pesquisa que está crescendo rapidamente, e muitos pesquisadores estão trabalhando para desenvolver novas tecnologias e aplicativos. A IA também está sendo usada em muitas indústrias, como finanças, saúde, educação e transporte. A IA é uma área de pesquisa que está crescendo rapidamente, e muitos pesquisadores estão trabalhando para desenvolver novas tecnologias e aplicativos.



O que é a vida



A vida é uma experiência humana que começa com o nascimento e termina com a morte. A vida é uma série de eventos que acontecem entre esses dois momentos. A vida é uma coisa que acontece com você, não algo que você faz. A vida é uma coisa que acontece com você, não algo que você faz. A vida é uma coisa que acontece com você, não algo que você faz. A vida é uma coisa que acontece com você, não algo que você faz. A vida é uma coisa que acontece com você, não algo que você faz.



Obrigado!

