

Notebook – Classificação de resenhas de livros

Leandro Carísio

Aplicação

Pare esse exercício, usei a base de dados ReLi [1].

São 1600 resenhas (em português) de livros de 7 autores.

O objetivo é classificar cada resenha em positiva, negativa ou neutra indicando se o leitor gostou ou não do livro.

Foi usada a API da OpenAI (gpt-3.5-turbo).

Conceitos importantes do exercício feito

Zero-Shot e Few-Shot

Se refere a quantidade de exemplos que são fornecidos no tempo de inferência ao apresentar uma tarefa para o modelo.

No zero-shot, apenas a tarefa é apresentada. No few-shot, além da tarefa, alguns exemplos são fornecidos.

Resultados – 1) classificação manual

A base de dados é anotada, mas não no escopo da resenha. Apenas de frases das resenhas. Assim, uma resenha pode ter frases anotadas com polaridade positiva ou negativa.

Por isso, foi feita a classificação manual de 100 resenhas escolhidas aleatoriamente pelo computador. As resenhas tem a seguinte distribuição:

```
Porcentagem por classificação:  
Positiva      0.68  
Neutra        0.18  
Negativa      0.14  
Name: classificacao_manual, dtype: float64
```

Nota: Há uma dose de subjetividade na classificação manual. Dependendo da resenha, outro classificador humano poderia escolher uma classificação diferente da minha.

Resultados – 2) Zero-shot - classificação

Formato da mensagem enviada para a API:

```
Responda em uma palavra se a resenha é positiva, negativa ou neutra:
```

```
Resenha:
```

```
###
```

```
AQUI VAI O TEXTO DA RESENHA DO LEITOR
```

```
###
```

Foi feita a classificação das 1.600 resenhas. A distribuição foi:

```
Porcentagem por classificação:
```

```
Positiva      0.716875
```

```
Negativa      0.163125
```

```
Neutra        0.120000
```

```
Name: classificacao_gpt, dtype: float64
```

```
Porcentagem por classificação:
```

```
Positiva      0.68
```

```
Neutra        0.18
```

```
Negativa      0.14
```

```
Name: classificacao_manual, dtype: float64
```

A porcentagem por tipo de classificação do GPT é relativamente parecida da classificação manual

Resultados – 2) Zero-shot – matriz de confusão

Matriz de confusão (classificação manual *versus* classificação automática, 100 exemplos):

classificacao_gpt	Negativa	Neutra	Positiva
classificacao_manual			
Negativa	12	2	0
Neutra	7	6	5
Positiva	2	5	61

A maior parte do erro ficou limitada a uma diferença de 1 classe. Houve apenas 2 exemplos, que o GPT classificou como negativo e a classificação manual foi positiva.

Obs.: A classificação manual foi feita antes de olhar os resultados da automática.

Resultados – 2) Zero-Shot – tipos de erros

Algumas resenhas são confusas e poderiam ter tido uma anotação diferente por outra pessoa:

Considere a resenha 1518 (manual: Negativa, GPT: Neutra):

1518: O livro é envolvente e tal , te prende em o começo , mas de o meio pro final parece meio forçado , o autor meio que começa a inventar histórias demais , não havendo uma ligação coerente entre elas , fica uma impressão de que se está lendo algum tipo de livro seriado , sei lá , com várias histórias diferentes em o mesmo livro . Não é a melhor obra de o grande Sidney Sheldon , mas recomendo , uma boa leitura pra passar o tempo .

163: Esse livro estou trocando .

Algumas apenas resenhas relatam a história e foram classificadas assim mesmo:

Resultados – 2) Zero-Shot – tipos de erros

Algumas resenhas contam coisas positivas/negativas sobre outras coisas que não o livro:

```
1354: nhaaa , cheguei a ver o filme nocinema , e nossa fiquei com mta vontade ler ... por  
isso vou ler mesmo .  hauhua
```

Dificuldade de identificar o pensamento dominante em caso de conflito de ideias

```
255: achei interessante .. mais prefiro o filme !
```


Resultados – 3) Few-shot - mensagem

Formato da mensagem:

Responda em uma palavra se a resenha é positiva, negativa ou neutra:

Resenha 1: Adorei como ele traz assuntos mais complicados e complexos de forma mais leve e simples, o que ajuda muito na hora do entendimento! Literalmente incrível, li em três dias e não parava de pensar no livro, esperava outro final mas e sse ficou de ouro também.

Resposta: Positiva

##

Resenha 2: Eu achei uma história fraca, um tanto sem amarração. Não consegui passar da metade. Sei que muitos gostam desse livro, mas eu achei uma ideia muito boa sem conclusão. Pode ser culpa da tradução. Não curti.

Resposta: Negativa

##

Resenha 3: Em A Biblioteca da Meia-Noite, Nora Seed se vê exatamente na situação pela qual todos gostaríamos de poder passar: voltar no tempo e desfazer algo de que nos arrependemos.

Resposta: Neutra

##

Resenha 4: AQUI VAI O TEXTO DA RESENHA DO LEITOR

Resposta:

Optei por pegar uma resenha positiva e uma negativa da Amazon sobre outro livro (A biblioteca da meia-noite). Para o exemplo neutro, utilizei parte da sinopse do livro

Resultados – 3) Few-shot – matriz de confusão

Matriz de confusão (classificação manual *versus* classificação automática, 100 exemplos):

classificacao_gpt		Negativa	Neutra	Positiva
classificacao_manual				
	Negativa	11	3	0
	Neutra	5	11	2
	Positiva	6	14	48

Muito mais erros do que no zero-shot (30 vs 21).

Resultados – 3) Few-shot – tipos de erros

Os tipos de erros são semelhantes do no zero-shot. Entretanto, nesse caso ocorreu alguns resultados inesperados:

1. Ao contrário do solicitado, às vezes o modelo informava o motivo da classificação:

“Neutra (não é uma resenha, mas sim um resumo da trama do livro)”

“Negativa (devido aos erros de ortografia e gramática)”

2. Mesmo “entendendo” que a resenha só conta a história, dava uma classificação:

“Negativa (a resenha não fala especificamente sobre o livro “Capitães de Areia”, mas sim sobre a história em si e a realidade social que ela retrata)”

Resultados – 4) Zero-shot vs Few-shot

Para esse exemplo, zero-shot teve resultados muito melhores que few-shot.

Uma possível explicação pode ser que, ao apresentar exemplos para o modelo, posso tê-lo induzido a exemplos de erros que vi ocorrendo no zero-shot fazendo com que o modelo “olhe” para determinados tipos de dados/padrões e deixe de generalizar tão bem quanto antes.

Obrigado

Leandro Carísio
carisio@gmail.com