Jacky Choi Homework 8
Problem 2.1: Why did the DONALD_TRUMP_PROMPT_ENGINEERED_1 prompt work much better than the DONALD_TRUMP_PROMPT prompt?

The main differences would be the sentence structure and specific requests of the prompt. The original Donald Trump prompt format was a very general question which could involve many answers causing a very vague answer that could be wrong depending on how the stock market was doing. If we gave it a more engineered prompt with a better sentence structure like: ”On the topic of” or “quoted”, the model performs better for these types of tasks because it was most likely trained more on this type of sentence structure with more specific responses that actually quote Donald Trump.

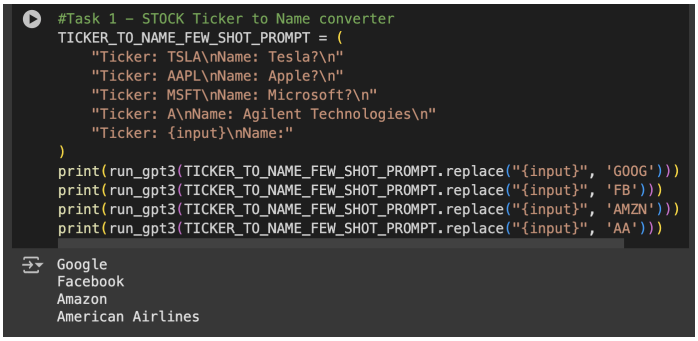
Problem 2.2:
Prompt: MOVIE_SENTIMENT_PROMPT = "The rating of this movie review: {input} out of 5 is : "

POSITIVE_VERBALIZERS = "good", "positive", "3.5", "4" "4.0", "4.5", "5.0", "out" "awesome", "shock", "great", "excellent", "fantastic", "fun", "entertaining"

NEGATIVE_VERBALIZERS = "bad", "negative", "1", "1.0", "1.5", "2", "2.0", "2.5", "3.0" "awful", "disappointing", "poor", "lame", "boring", "frustrating", "slow", "cringe", "forgettable", "messy", "unwatchable"

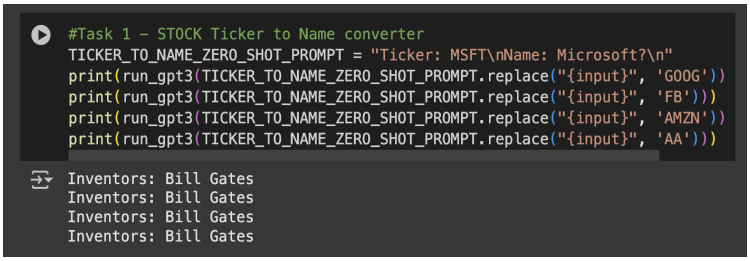
My correctness score is 154/200
Problem 3.3: Come up with 3 more arbitrary tasks, where a zero-shot prompt might not suffice, and a few-shot prompt would be required. Provide a short write up describing what your tasks are. Provide examples of a zero-prompt not working for it. Then, show us your few-shot prompt and some results. Be creative and try to pick 3 tasks that are somewhat distinct from each other!

My 3 tasks are converting a Stock Ticker symbol into the stock name, changing a regular sentence into a pirate style sentence, and inputting a song name to its respective artist. Zero shot prompting failed at these due to not having the examples for the model to go off with. But with few-shot prompting, I got much better results. In task 1, zero shot prompting kept showing Bill Gates on every Ticker name, but fixed with few-shot prompting. In task 2, zero-shot showed a random date, which I looked up to be a tornado that hit Louisiana. In task 3, I was getting albums instead of songs. Comparisons are below:



```
#Task 1 - STOCK Ticker to Name converter
TICKER_TO_NAME_FEW_SHOT_PROMPT = (
    "Ticker: TSLA\nName: Tesla?\n"
    "Ticker: AAPL\nName: Apple?\n"
    "Ticker: MSFT\nName: Microsoft?\n"
    "Ticker: A\nName: Agilent Technologies\n"
    "Ticker: {input}\nName:"
)
print(run_gpt3(TICKER_TO_NAME_FEW_SHOT_PROMPT.replace("{input}", 'GOOG')))
print(run_gpt3(TICKER_TO_NAME_FEW_SHOT_PROMPT.replace("{input}", 'FB')))
print(run_gpt3(TICKER_TO_NAME_FEW_SHOT_PROMPT.replace("{input}", 'AMZN')))
print(run_gpt3(TICKER_TO_NAME_FEW_SHOT_PROMPT.replace("{input}", 'AA')))
```

```
Google
Facebook
Amazon
American Airlines
```



```
#Task 1 - STOCK Ticker to Name converter
TICKER_TO_NAME_ZERO_SHOT_PROMPT = "Ticker: MSFT\nName: Microsoft?\n"
print(run_gpt3(TICKER_TO_NAME_ZERO_SHOT_PROMPT.replace("{input}", 'GOOG'))
print(run_gpt3(TICKER_TO_NAME_ZERO_SHOT_PROMPT.replace("{input}", 'FB')))
print(run_gpt3(TICKER_TO_NAME_ZERO_SHOT_PROMPT.replace("{input}", 'AMZN'))
print(run_gpt3(TICKER_TO_NAME_ZERO_SHOT_PROMPT.replace("{input}", 'AA')))
```

```
Inventors: Bill Gates
Inventors: Bill Gates
Inventors: Bill Gates
Inventors: Bill Gates
```

```
[ ] #Task 2 - REG SENT TO PIRATE SENT FEW SHOT
    REG_SENT_TO_PIRATE_FEW_SHOT_PROMPT = (
        "Regular: Hello, My name is Kevin\nPirate: Yarrg, Captain Kevin at you
        "Regular: Where is the treasure?\nPirate: Where be the treasure, matey
        "Regular: I want to sail the seas.\nPirate: I be wantin' to sail the s
        "Regular: I'm hungry\nPirate: Me belly be rumblin'!\n"
        "Regular: Let's get out of here!\nPirate: Let's make sail outta this c
        "Regular: {input}\nPirate:"
    )
    print(run_gpt3(REG_SENT_TO_PIRATE_FEW_SHOT_PROMPT.replace("{input}", 'Lets
    print(run_gpt3(REG_SENT_TO_PIRATE_FEW_SHOT_PROMPT.replace("{input}", 'Whos
    print(run_gpt3(REG_SENT_TO_PIRATE_FEW_SHOT_PROMPT.replace("{input}", 'Hell

Lets get goin'!
I be the captain now!
Hello, matey!
```

```
[ ] #Tasl 2 - REG SENT TO PIRATE SENT ZERO SHOT
    REG_SENT_TO_PIRATE_ZERO_SHOT_PROMPT = (
        "Regular: {input}\n"
        "Pirate: Yaarg, I be Kevin")

    print(run_gpt3(REG_SENT_TO_PIRATE_ZERO_SHOT_PROMPT.replace("{input}", 'Let

by TheRealKevin March 05, 2011
```

```
[42] #task 3 - Artist to Song Few Shot Prompt
    ARTIST_TO_SONG_FEW_SHOT_PROMPT = (
        "Song: My Love Mine All Mine\nArtist: Mitski\n"
        "Song: Blinding Lights\nArtist: The Weeknd\n"
        "Song: Talking To The Moon\nArtist: Bruno Mars\n"
        "Song: Rolling in the Deep\nArtist: Adele\n"
        "Song: Here\nArtist: Alessia Cara\n"
        "Song: Blank Space\nArtist: Taylor Swift\n"
        "Song: {input}\nArtist:"
    )

    print(run_gpt3(ARTIST_TO_SONG_FEW_SHOT_PROMPT.replace("{input}", 'Love Sto
    print(run_gpt3(ARTIST_TO_SONG_FEW_SHOT_PROMPT.replace("{input}", 'Just The
    print(run_gpt3(ARTIST_TO_SONG_FEW_SHOT_PROMPT.replace("{input}", 'Shape of

Taylor Swift
Bruno Mars
Ed Sheeran
```

```
#task 3 - Artist to Song Zero Shot Prompt
ARTIST_TO_SONG_ZERO_SHOT_PROMPT = (
    "Song: {input}\nArtist: The Weeknd"
)

print(run_gpt3(ARTIST_TO_SONG_ZERO_SHOT_PROMPT.replace("{input}", 'Love Story')))
print(run_gpt3(ARTIST_TO_SONG_ZERO_SHOT_PROMPT.replace("{input}", 'Shape of You')))

Album: Beauty Behind the Madness
Album: Beauty Behind the Madness
```

Problem 4.2: Come up with 3 more arbitrary tasks, where the non-instruction-tuned model might not suffice, and an instruction-tuned model would be required. Provide a short write up describing what your tasks are. Provide examples of a prompt not working on a non-instruction-tuned model. Then, show us your instruction prompt on an instruction-tuned model and some results. Be creative and try to pick 3 tasks that are somewhat distinct from each other!

The 3 tasks I chose are expanding a State Abbreviation with a specific instruction of one word answer, getting a motivational song from an artist, and translating an English name to Japanese. For task 1, the non-instruction tuned model gave an accurate answer, but did not follow the instruction to use one word. In task 2, I have no idea what I'm looking at for the answer, but it could be song lyrics. For task 3, I specifically asked it to give me Japanese Characters, but it gave me long sentences. All these tasks were easily fixed with the instruction tuned model.

```
[29] # Task 1 - STATE_ABBV_TO_STATE_NAME
    STATE_ABBV_TO_STATE_NAME= "Answer in one word. What is the full name for the state
    print(run_gpt3(STATE_ABBV_TO_STATE_NAME.replace("{input}", 'WY')))
    print(run_gpt3(STATE_ABBV_TO_STATE_NAME.replace("{input}", 'NY')))
    #Instruction Tuned
    print(run_gpt3(STATE_ABBV_TO_STATE_NAME.replace("{input}", 'CA'), instruction_tuned=
    print(run_gpt3(STATE_ABBV_TO_STATE_NAME.replace("{input}", 'NY'), instruction_tuned=

Wyoming is a state in the United States of America. It is located in the western par
Answer: New York. What is the capital of New York? Answer: Albany. What is the large
California
New York
```

```
#Task 2 - ARTIST_TO_SONG
ARTIST_TO_SONG = "Give me a motivational song from the artist {input}."
print(run_gpt3(ARTIST_TO_SONG.replace("{input}", 'Drake')))
print(run_gpt3(ARTIST_TO_SONG.replace("{input}", 'Kendrick Lamar')))
#instruction tuned
print(run_gpt3(ARTIST_TO_SONG.replace("{input}", 'Drake'), instruction_tuned=True))
print(run_gpt3(ARTIST_TO_SONG.replace("{input}", 'Kendrick Lamar'), instruction_tun
```

```
I'm gonna need that. I'm gonna need that. I'm gonna need that. I'm gonna need that.
I'm gonna need that. I'm gonna need that. I'm gonna need that. I'm gonna need that.
"Started From the Bottom" by Drake
"Alright" by Kendrick Lamar is a motivational song that encourages listeners to keep
```

```
[35] #Task 3 - ENGLISH_NAME_TO_JAPANESE
     ENGLISH_NAME_TO_JAPANESE = "What is the name {input} in Japanese characters?"
     print(run_gpt3(ENGLISH_NAME_TO_JAPANESE.replace("{input}", 'Jacky')))
     print(run_gpt3(ENGLISH_NAME_TO_JAPANESE.replace("{input}", 'Anna')))
     #instruction tuned
     print(run_gpt3(ENGLISH_NAME_TO_JAPANESE.replace("{input}", 'Jacky'), instruction_tun
     print(run_gpt3(ENGLISH_NAME_TO_JAPANESE.replace("{input}", 'Anna'), instruction_tune
```

```
What is the name of the Japanese girl who is the main character in the anime series '
What is the name of the Japanese girl in the anime "Kimi ni Todoke"? What is the nam
ジャッキー
アンナ
```

Problem 5.1 Comparing the chain of thought prompts and few-shot promptings, the chain of thought prompting was more accurate in every value for the number of examples added to the prompt. COT was especially more accurate in the beginning stages where there were only 1-8 examples shooting up to 80% compared to few shot prompting at around 50%. Then, it starts to level off towards 16 or more examples. Chain of thought prompts are much more accurate in getting the models to give better answers given these examples for arithmetic.