SEP 775 - Assignment 4

Final Project

Prepared by:

Sizhe Guan	400184473
Haolan Liu	400486179
Luke Vanden Broek	400486889

April 28, 2023

Problem Statement 2
Paper Reproduction 2
- **Literature Review** 2
- **Model Approach and Implementation** 3
- **Results** 8

Novel Implementation 12
- **Dataset Creation** 12
- **Model Approach and Implementation** 14
- **Results** 15

Evaluation of Models 19
Conclusion 19
References 21

Problem Statement

In this report, we present our findings on further fine-tuning the GPT-2 model from our previous assignment. We also include the findings of our previous report, and the model we recreated in that assignment. In the previous assignment, a literature review was performed to investigate the topic of NPC dialogue systems for role-playing video games (RPGs). We then fine-tuned the base GPT-2 model on a large dataset of Word of Warcraft (WOW) quests based off of the work done in another paper (van Stegeren & Myśliwiec, 2021). This allowed for the automated generation of quest titles, objectives and descriptions, as well as generating quest descriptions based off of the quest title and objective.

In this project, we take the fine-tuned model created in the previous assignment and further fine-tune it on a new dataset we created based off of the Star Wars: The Old Republic (SWTOR) quest database (*Quests,* n.d.a). The goal was to update the model to use the same quest structure, but in a way that quests are created that fit thematically within the word of SWTOR rather than WOW.

Paper Reproduction

Literature Review

Previous works on NPC dialogue generation can be categorized into two major categories: dialogue tools and natural language processing-based technologies (Kerr & Szafron, 2009). Compared to the dialogue tools approach, natural language processing-based technologies reduced the work on dialogue creation. Kerr and Szafron (Kerr & Szafron, 2009) discussed the advantage of automated dialogue generation in video game production and how NLP technologies can be used to reduce the workload for game development and speed up the development process. However, dialogue tools are the major choice as dialogue generated by natural language processing technologies is simplistic and rigidly structured which could influence the player experience and reduce the controllability of conversations from game designers.

One representation of an NLP-based NPC dialogue system is the dialogue generation system used in Façade (Mateas & Stern 2003). The game used a custom NLU Template Language to process the text from users and predict potential reactions, references to objects and related topics based on customized local context. In 2016, Kacmarcik used tree regular expression (trex) patterns to identify common structures, extract the interaction pairs for NPC dialogue, and generate possible responses by producing all combinations of words based on tree manipulations (Kacmarcik, 2006). Other natural language processing approaches being used to generate or help NPC dialogue content include procedure content generation (PCG) (Hämäläinen, 2022), LSTM (Dai, 2020), AffectON (Bucinca et al., 2020), GPT-2 (Kalbiyev, 2022) (Ashby et al., 2023), knowledge graph (Ashby et al., 2023), and other language models.

In RPG games, side-quests are also a major component of a game to increase the game content and provide players with the freedom to explore the content. Different from major

quests, contents of side-quests usually follow a certain fixed structure (van Stegeren & Myśliwiec, 2021) and those contents are decorative text which would not influence the overall gameplay. Dialogue tools and systems had been developed and used for the generation for side-quests dialogue content, however, van Stegeren and Myśliwiec noticed that the application of transformer - based natural language processing technologies on NPC dialogue based on quests had not been discussed. GPT series is a series of powerful pretrained transformer neural network architecture language models published by OpenAI. State-of-arts GPT models, including GPT-3 (Brown et al., 2020) and GPT-4 (OpenAI, 2023), are not public resources. Based on current public resources, GPT-2 (Alec et al., 2019) is the most effective and powerful GPT model. GPT 2 was published in 2019 with the paper *Language Models are Unsupervised Multitask Learners* (Alec et al., 2019). The model has 1.5 billion parameters, trained on 8 million web pages in multiple domains curated or filtered by humans (Better Language Models and Their Implications, n.d.) in a self-supervised approach. The training task is to predict the next token for a given content. Concluded by OpenAI, the model is able to perform well across various domains and datasets, and can even exceed the performance of some domain-specified models on certain tasks without needing to use domain-specific training datasets.

The dataset required for this project is text containing NPC dialogue, quest titles, or objectives to provide context for the NPC dialogue. To obtain this dataset, we have estimated that using data from Massive Multiple Online RPGS (MMORPGs) would be ideal as they tend to have more playable content than offline or single-player games. Additionally, these games usually have large player communities, which could mean that quest information is readily available online, for instance, on fan websites. By using this data source, we can collect a large amount of text data necessary for training and testing our model accurately.

Model Approach and Implementation

The paper the team chose to implement was on “Fine-tuning GPT-2 on annotated RPG quests for NPC dialogue generation” (van Stegeren & Myśliwiec, 2021). This paper investigates working with the GPT-2 model available on Google Collaboratory and fine-tuning it on a dataset containing in-game quests from the massively multiplayer online role-playing game (MMORPG) World of Warcraft (WOW) (van Stegeren & Myśliwiec, 2021). The goal of this paper was to create a model that could generate interesting descriptions for in-game quests that players could undergo that would fit within the lore of the game (van Stegeren & Myśliwiec, 2021).

The dataset they used was created using a quest database containing important information about the quests (*Quests,* n.d.a). They then annotated the dataset to focus on three key pieces of information: the quest title, objective and description, as seen in the sample in Figure 1. The dataset was made available to the public, so we used their provided dataset rather than creating and annotating the dataset again due to time constraints.

Structure
<|startoftext|> [quest title] <|obj|> [quest objective] <|text|> [quest description] <|endoftext|>
Example datapoint
<|startoftext|> In Dire Need <|obj|> Hear out the Council of Six in the Purple Parlor. <|text|> The ill tidings you bear only increase the concerns the Council has been having. I know you have your hands full with the Tomb of Sargeras. Make no mistake, the battle for the tomb remains our top priority–but we must not think the Legion foolish enough to rely on brute force alone. They are far more cunning than that. Please, hear us out. <|endoftext|>

Figure 1: Training data structure and sample in dataset

In order to re-implement the work done in the paper, the GPT-2 model was loaded and trained based off of a gpt-simple tutorial (Woolf, 2019). Firstly, the model was imported, and the GPU used was confirmed to be the NVIDIA Tesla T4 GPU provided by Google Collaboratory, as observed in Figure 2.

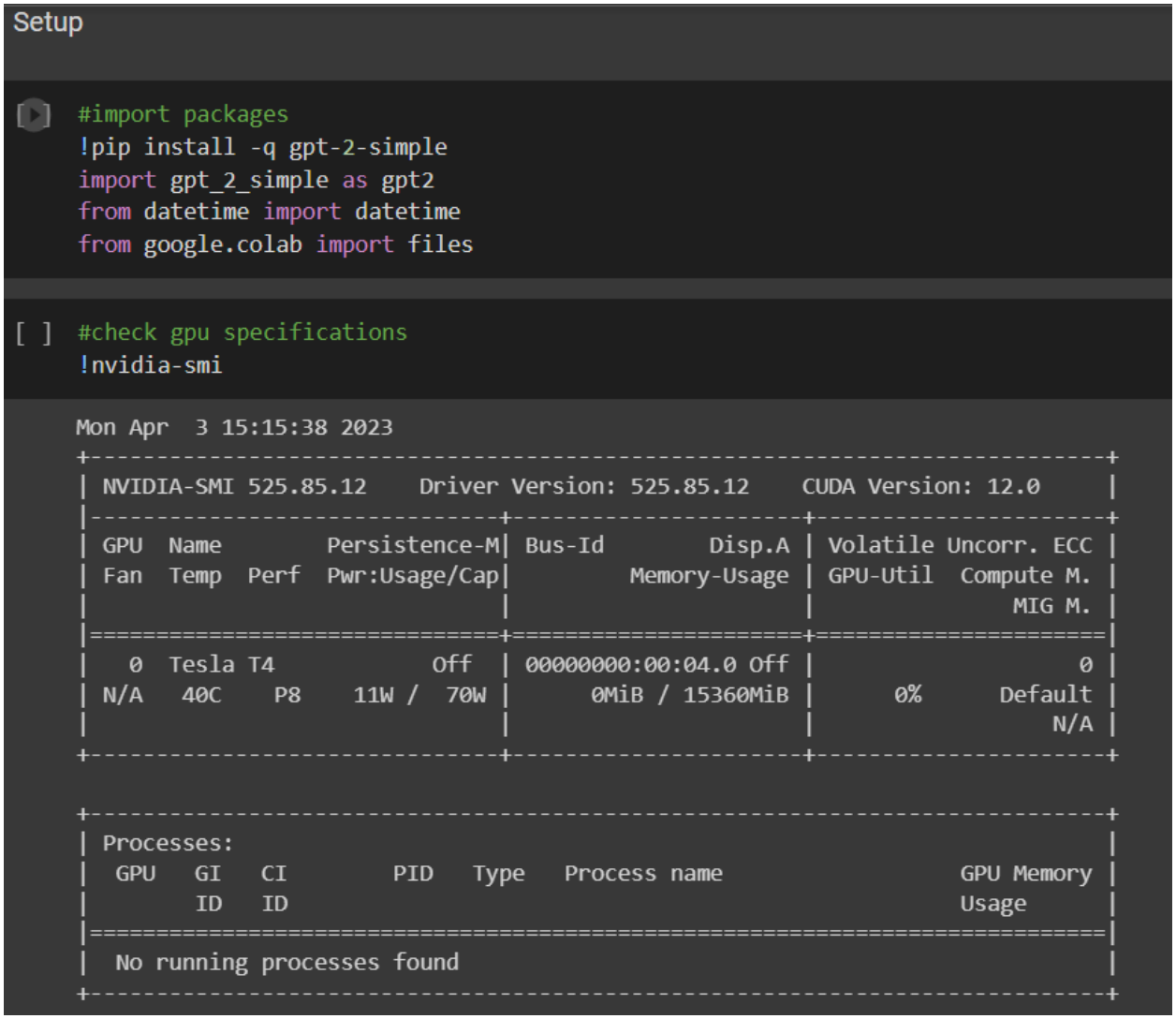


Figure 2: Importing packages and GPU specifications

The 124M version of the model was then downloaded for use,and the Collaboratory file was mounted in Google Drive as seen in Figure 3, and the dataset from the original paper was downloaded as observed in Figure 4.

```
#download model
gpt2.download_gpt2(model_name="124M")
```

```
Fetching checkpoint: 1.05Mit [00:00, 294Mit/s]
Fetching encoder.json: 1.05Mit [00:00, 2.75Mit/s]
Fetching hparams.json: 1.05Mit [00:00, 734Mit/s]
Fetching model.ckpt.data-00000-of-00001: 498Mit [00:16, 30.8Mit/s]
Fetching model.ckpt.index: 1.05Mit [00:00, 635Mit/s]
Fetching model.ckpt.meta: 1.05Mit [00:00, 3.69Mit/s]
Fetching vocab.bpe: 1.05Mit [00:00, 3.69Mit/s]
```

```
#mount google drive
gpt2.mount_gdrive()
```

```
Mounted at /content/drive
```

Figure 3: Downloading the model and mounting google drive

```
#load WOW quest dataset
!wget jakub.thebias.nl/GPT2_WOWHead_dataset.txt
```

```
--2023-04-03 15:18:38--  http://jakub.thebias.nl/GPT2_WOWHead_dataset.txt
Resolving jakub.thebias.nl (jakub.thebias.nl)... 130.89.161.150, 2001:67c:2564:331:6ab5:99ff:fe6d:772c
Connecting to jakub.thebias.nl (jakub.thebias.nl)|130.89.161.150|:80... connected.
HTTP request sent, awaiting response... 301 Moved Permanently
Location: https://jakub.thebias.nl/GPT2_WOWHead_dataset.txt [following]
--2023-04-03 15:18:38--  https://jakub.thebias.nl/GPT2_WOWHead_dataset.txt
Connecting to jakub.thebias.nl (jakub.thebias.nl)|130.89.161.150|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 11704501 (11M) [text/plain]
Saving to: ‘GPT2_WOWHead_dataset.txt’

GPT2_WOWHead_datase 100%[===================>]  11.16M  9.06MB/s    in 1.2s

2023-04-03 15:18:40 (9.06 MB/s) - ‘GPT2_WOWHead_dataset.txt’ saved [11704501/11704501]
```

Figure 4: Loading the dataset

The GPT model was then fine-tuned, as seen in Figure 5. As noted earlier, the dataset is loaded and the model used is the 124M model in our experiment. The model ran for 1000 steps. We set the model to restore from ‘fresh’, meaning we train from the base GPT-2 model. The model is also set to print training progress every 10 steps, and print an example output every 500 steps and save progress every 500 steps (halfway through).

```
sess = gpt2.start_tf_sess()

gpt2.finetune(sess,
              dataset="GPT2_WOWHead_dataset.txt",
              model_name='124M',
              steps=1000,
              restore_from='fresh',
              run_name='run1',
              print_every=10,
              sample_every=200,
              save_every=500
              )
```

```
Loading checkpoint models/124M/model.ckpt
Loading dataset...
100%|██████████| 1/1 [00:12<00:00, 12.63s/it]
dataset has 3119753 tokens
Training...
[10 | 29.79] loss=2.71 avg=2.71
[20 | 51.72] loss=2.72 avg=2.71
[30 | 74.04] loss=2.23 avg=2.55
[40 | 96.78] loss=2.81 avg=2.62
[50 | 119.92] loss=2.62 avg=2.62
[60 | 143.50] loss=2.63 avg=2.62
[70 | 167.31] loss=2.55 avg=2.61
[80 | 190.79] loss=2.59 avg=2.61
[90 | 214.27] loss=2.65 avg=2.61
[100 | 237.93] loss=2.56 avg=2.61
[110 | 261.53] loss=2.46 avg=2.59
[120 | 285.08] loss=2.61 avg=2.59
[130 | 308.63] loss=2.70 avg=2.60
[140 | 332.26] loss=2.69 avg=2.61
```

Figure 5: Training code and output of first 140 steps

After this, the model checkpoint could be saved and loaded for future use as seen in Figure 6. This allows for the model to be tested and used on new data, as well as allowing for further fine-tuning in the future.

```
[ ]  #save dataset checkpoint to google drive
     gpt2.copy_checkpoint_to_gdrive(run_name='run1')

[ ]  # load the model from google drive
     gpt2.copy_checkpoint_from_gdrive(run_name='run1')
```

Figure 6: Saving and loading the checkpoint

The following code could then be run after restarting the runtime to begin observing results.

```
#restart runtime after previous blocks, then run this block
!pip install -q gpt-2-simple
import gpt_2_simple as gpt2
from datetime import datetime
from google.colab import files

sess = gpt2.start_tf_sess()
gpt2.load_gpt2(sess, run_name='run1')
```

Figure 7: Code to start testing the model

Results

After training the model and reloading it into google colab, it could now be tested on data. Firstly, an output can be observed below to generate some random texts with quest names, objectives and text descriptions using the code in Figure 8 to produce the following outputs.

```
gpt2.generate(sess, run_name='run1')
```

Figure 8: Random quest generation results

<|endoftext|>
<|startoftext|>The Headless Horseman<|obj|>See if you can recover the head of the Headless Horseman.<|text|>The head of the Headless Horseman has been stolen. It's too late for him now, but it's not too late for him now for that head to stop falling off the horse. You'll find the horseman sitting on the floor of the cave to the east, so that's where you can get your head. There's a large stone pillar that you can focus on, and I can slow it down. I'll need you to

recover it, and I'll need you to do it fast indeed!<|endoftext|>
<|startoftext|>Staking Out the Archmage<|obj|>Speak to Archmage Pamela at the Shrine of the Crimson Tide in Dalaran.<|text|>The Archmage has ordered you to report to her at the Shrine of the Crimson Tide. Our goal is to strike at the heart of the Crimson Tide and take its power.
She will want to hear your words.<|endoftext|>
<|startoftext|>Bounty on the Headless Horseman<|obj|>Collect the Headless Horseman's Head from a Shadow Knight in the Abandoned Ruins in the Felwood Hills.<|text|>The headless horseman's head is a great deal more dangerous than he thought, but he is the head of a great lie. He is a powerful that will enslave your people, you and I. The scoundrel is a powerful weapon in your hands.
We need to find the head of the headless horseman. I will need your help.<|endoftext|>
<|startoftext|>The Headless Horseman<|obj|>Meet with the Archmage Pamela at the Shrine of the Crimson Tide in Dalaran.<|text|>The Headless Horseman has been taken prisoner. He's a powerful that will enslave your people, you and I.
We need to get him out of there. Meet with the Archmage and ask him if he's alright.<|endoftext|>
<|startoftext|>A Feast for the Headless Horseman<|obj|>Meet with the Archmage Pamela at the Shrine of the Crimson Tide in Dalaran.<|text|>The headless horseman's head is a great deal more dangerous than he thought, but he is the head of a great lie. He is a powerful that will enslave your people, you and I.
We need to get him out of there. Meet with the Archmage and ask him if he's alright.<|endoftext|>
<|startoftext|>A Feast for the Headless Horseman<|obj|>Meet with the Archmage Pamela at the Shrine of the Crimson Tide in Dalaran.<|text|>The headless horseman's head is a great deal more dangerous than he thought, but he is the head of a great lie. He is a powerful that will enslave your people, you and I.
We need to get him out of there. Meet with the Archmage and ask him if he's alright.<|endoftext|>
<|startoftext|>The Headless Horseman<|obj|>Meet with Archmage Pamela at the Shrine of the Crimson Tide in Dalaran.<|text|>The headless horseman's head is a great deal more dangerous than he thought, but he is the head of a great lie. He is a powerful that will enslave your people, you and I.
We need to get him out of there. Meet with the Archmage and ask him if he's alright.<|endoftext|>
<|startoftext|>The Headless Horseman<|obj|>Meet with Archmage Pamela at the Shrine of the Crimson Tide in Dalaran.<|text|>The headless horseman's head is a great deal more dangerous than he thought, but he is the head of a great lie. He is a powerful that will enslave your people, you and I.
We need to get him out of there. Meet with the Archmage and ask him if he's alright.<|endoftext|>
<|startoftext|>The Headless Horseman<|obj|>Meet with Archmage Pamela at the Shrine of the Crimson Tide in Dalaran.<|text|>The headless horseman's head is a great deal more dangerous than he thought, but he is the head of a great lie. He is a powerful that will enslave your people, you and I.
We need to get him out of there. Meet with the Archmage and ask him if he's alright.<|end

Quest dialogue can be generated for a specific quest name and quest objective as well, as seen in Figure 9. In this example, we set the maximum token length to 250 and the

temperature to the default value of 0.7. Higher temperatures can lead to more unexpected outputs of the model, which can be more interesting and surprising, but could also lead to less coherent results (van Stegeren & Myśliwiec, 2021). The prefix contains the prompt we use to generate the description, and we use truncate to break the text whenever the <|end of text|> statement is encountered.

```
#generated text results for specifc quest name and objective
# format <|startoftext|> Quest Name <|obj|> Quest Objective <|text|>
gpt2.generate(sess,
              length=250,
              temperature=0.7,
              prefix="<|startoftext|>Spooky Cave Party<|obj|>Talk to the goblins in the Big Spooky Cave.<|text|>",
              truncate="<|endoftext|>",
              nsamples=5,
              batch_size=5,
              include_prefix=False,
              )
```

Figure 9: Generating text for a specific combination prompt

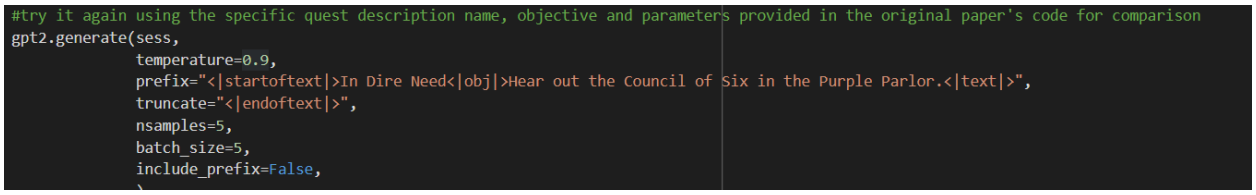
In this case, the quest name input into the system is “Spooky Cave Party”, with a quest objective of “Talk to the goblins in the Big Spooky Cave.” Five sample outputs can be seen below.

Goblins! They're out there, exploring the cave! I don't mind if I'm just a little kid, but I'm really, really worried about these goblins! I don't know what to do! Oh, I need a little help. You can't go crazy trying to go crazy! If you were to go crazy, I'd be really... excited about it! You go talk to the goblins in the Big Spooky Cave, and if they're up for something, maybe they'll give you a ride!
==========
Hey, dang goblin! Friend of mine, you're not going to get in to a cave party without getting a treat. The goblins are making fun of us, and I'm pretty upset that I didn't see what they were doing earlier. I dunno what they're up to, but I'm sure they'll be fine. Let's see if you can bring some friends and go out to the Big Spooky Cave. Just stop in the middle of the cave to talk to the goblin in charge.
==========
Ricky's horrible story. He was a goblin! He was in the big cave during the show when he went to the back of the cave and he was crawling onto my little rumpus. I thought he was going to die, but what the heck did he do?
I saw him crawling on the floor, and I was so mad! I said, "Ricky, you're gonna have to go back into the cave and deal with this man!"
I was so mad! And then he came back and said, "Ricky, you're gonna have to go back to the cave and deal with this guy."
I said, "Ooooh! Ricky your dead!"
I heard that he was all over the place. Then he came back and his whole cave exploded! How did you survive that show?
==========
Have you heard about the goblins in the cave? They're the only goblins in the cave. If you want to help out the goblins in any way, you'll have to go talk to the goblins. Besides, I wouldn't mind helping out along the way.
==========

The dwarves are preparing the cave for a scary cave party, so I'm going to go ahead and put the goblins to work. There's some goblin-looking dwarves in there, too! Go talk to them. They'll help you out with the party.
==========

To observe how our version of the model compares with the original paper's, we generated text with the example prompt used in the original code, and used the same parameters for generation. We used the code in FIgure 10 to generate our results for the quest name "In Dire Need", and the quest objective "Hear out the Council of Six in the Purple Parlor." In this case, they used a temperature of 0.9, so we set ours to this value as well.

.



```
#try it again using the specific quest description name, objective and parameters provided in the original paper's code for comparison
gpt2.generate(sess,
              temperature=0.9,
              prefix="<|startoftext|>In Dire Need<|obj|>Hear out the Council of Six in the Purple Parlor.<|text|>",
              truncate="<|endoftext|>",
              nsamples=5,
              batch_size=5,
              include_prefix=False,
              )
```

Figure 10: Results for the combination prompt provided in the original paper's code

The results from the original paper's code can be seen below.

We're bound to survive, but we can't afford to put off the next battle indefinitely. The fight against the Burning Legion rages on within Ar'jora. We must hear the Council of Six's plans immediately. Don't wait to hear their concerns over their sleep. Speak to them when you are ready to hear the Council's plans.
==========
Well now, it's turn to Glory Hightide and Glory Hightide for a very simple task. I have written a letter to the Council of Six requesting that they send them some boots. I'm hoping its their response that will help us gain our protection. You proved yourself useful a few days ago, I'll reward you. None of us want to lose a good ally now do we? You can hear the Council of Six sitting at the purple parlor in the highest wing of the Wakening Council.
==========
Surely this means our captain is dead?
He was not the only trader to be taken captive. Most of our supplies were taken as well, but Rakjan and I aren't sure of the fate of them.
The captain's finance officer, the one I assumed to be Deepholm, is probably in the Purple Parlor. There's no telling how deep she went, but when she emerged, she was out of our reach.
==========
The message from the Council of Sages has not been relayed to Cabrinal. I'm worried sick. I am going to stay here with Illidan until something is done about this. Do you think you could go back to the Council and hear my worries? I don't like to imagine that someone is in danger. Or worse yet the Council is dead. I'd appreciate it if you would tell them I sent you.
==========
Darn horses and stuff! I said, it's a disaster! In case you haven't noticed, our mounts are in need of maintenance. Very badly in fact. Hurry to the purple parlor and tell 'em about the traffic jam the horses are causing this morning. Our top leaders would be eternally grateful.

```
==========
```

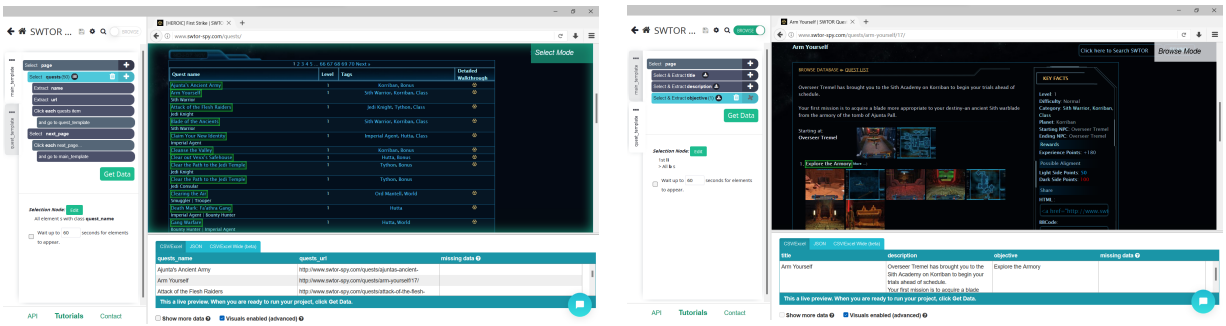
The next block showcases the results from our model using the same settings.

```
While our leadership team is here, its members are here working on a plan to return the gate to the green light. As you can see, it's not a well thought out system, so any team member who seeks to return to the gate shall be subject to discipline and commotion. They will be made to pay for the "incapacitated" actions of their production team.
==========
You have succeeded. You are truly our allies.
You have proven yourself to be a talented young mage and a valuable ally.
The Council of Six will have a direct exchange with you at once.
==========
Until recently, we took great pride in our reputation as champions. While it was our clansho to claim all four lands in this war; we're using those lands to test the dragons here - and our own. Havering to the north, you'll find Council of Six. Speak to them about your scouting of Un'Goro - it should inform your administration.
==========
I will begin your project of reconstruction. Present me proof of your progress so it may be your plan of attack. Here-- the reconstruction of our city. , witness what my men have to suffer in their attempt to wrest control of the city to the Scarlet Crusade and the Dark Iron dwarves. Having this Alliance from the Iron dwarves will not only release me from your control over them, it will also allow you to see the Legion's vision over our ruined city. When you're done, work to healing your city.
==========
It is time for you to go the Council of Six. Tell them that it is time to discuss the Alliance's plans in Darkshore. They can tell their soldiers the Alliance is ready for reinforcements at any time and from any enemy it comes into contact with.
==========
```

Novel Implementation

Dataset Creation

For this project, we aimed to fine-tune the GPT-2 model on the SWTOR quest database to observe if the model could be applicable to different games (*Quests,* n.d.a). In order to do this, it was necessary to create a dataset that fit the format of the WOW dataset with the content of the quests of the SWTOR dataset. To do this, we scraped the SWTOR database website using ParseHub, a tool that can be used to extract data from webpages (ParseHub, n.d.). ParseHub was able to go through each quest's web page and extract the description, objective and title from each of them as each quest webpage has the same format. The interface for ParseHub can be observed in Figure 11.



(a) (b)

Figure 11: Parsehub Interface (a) Sample page of quest list (b) Sample page of quest details

After processing through ParseHub, the data was exported to a .csv file containing 3471 quests. Prior to applying further preprocessing, we used Microsoft Excel to find and delete all quests that did not include an objective and/ or description, as they would not be useful in the final model. This resulted in a new dataset size of 2651 total quests, which can be observed in Figure 12.

	A	B	C	D	E	F	G	H	I
1	quests_titl	quests_de:	quests_objective						
2	Ajuntaâ€™	On your	Defeat Ancient Droids						
3	Arm Yours	Overseer	Explore the Armory						
4	Attack of t	Flesh	Speak to Master Satele Shan						
5	Claim Your	Your first	Speak to Jheeg						
6	Cleanse th	You have k	Kill Kâ€™lorâ€™slug Foragers						
7	Clear out \	Vexx has q	Eliminate the Hired Outlaws						
8	Clear the F	The	Defeat Flesh Raider Lurkers						
9	Clear the F	The	Defeat Flesh Raider Lurkers						
10	Clearing th	Sergeant	Disable Signal Jammer Alpha						
11	Death Mar	A bounty	Defeat Faâ€™athraâ€™s Gangsters						
12	Gang Warf	An ex-	Speak to Lew Brell						
13	Gauntlet	Braden, a	Search for Information on Vexx						
14	Halt the As	The cave t	Defeat Flesh Raiders						
15	Hit the Grc	Youâ€™r	Speak to Gearbox						
16	Kâ€™lorâ€	Sergeant	Kill Kâ€™lorâ€™slugs						
17	Landing Pa	After	Destroy the Remote Control Stations						
18	Leave No \	Thieves ha	Eliminate Faâ€™athraâ€™s Thieves						
19	Of Mind ar	You have	Enter the Sacrificial Chamber						
20	Prove You	Give Overs	Kill Kâ€™lorâ€™slug Cripplers						
21	Reclaiming	You have k	Kill Kâ€™lorâ€™slug Foragers						
22	Take Back	The separa	Dispatch the Separatist Forces						
23	The Path c	You have	Take the Speeder to the Gnarls						
24	A Fighting	You have	Speak to the Sith Inquisitor Trainer						

FIgure 12: .csv dataset file to be pre-processed

With the file ready to undergo further preprocessing, it was then imported into a new .ipynb file in Google Collaboratory for refinement. This code first removes any new lines from the quest descriptions, to keep each full quest in a single line for organization purposes. The code also removes any bracketed items from all titles, objectives and descriptions, as these usually refer to tags that might be relevant to the player, but are not relevant to the in-game lore

or story and therefore are not relevant to generating descriptions. These include tags such as [PvP], [HEROIC] or [INSTANCE]. Then, we removed any whitespaces from the beginning and end of each title, objective and description. Lastly, we converted from a .csv form to a .txt form with each item separated by the tags used in the WOW dataset. The code for this can be seen in Figure 13.

```
#cleaning data

for i in range(base_df.shape[0]):
  base_df["quests_description"][i]=base_df["quests_description"][i].replace('\n', ' ')#remove new lines from descriptions, replace with spaces

  #remove bracketed items (ex. [PVP], [repeatable], etc. (not relevant to context of story content))
  base_df["quests_title"][i] = re.sub("[\(\[].*?[\)\]]", "", base_df["quests_title"][i])
  base_df["quests_objective"][i] = re.sub("[\(\[].*?[\)\]]", "", base_df["quests_objective"][i])
  base_df["quests_description"][i] = re.sub("[\(\[].*?[\)\]]", "", base_df["quests_description"][i])

  #remove start and end whitespace
  base_df["quests_title"][i] = base_df["quests_title"][i].strip()
  base_df["quests_objective"][i] = base_df["quests_objective"][i].strip()
  base_df["quests_description"][i] = base_df["quests_description"][i].strip()

#converting format to format: "<|startoftext|>quests_title<|obj|>quests_objective<|text|>quests_description<|endoftext|>"
f = open("swtor_final_dataset.txt", "w")
for i in range(base_df.shape[0]):
  f.write("<|startoftext|>")
  f.write(base_df["quests_title"][i])
  f.write("<|obj|>")
  f.write(base_df["quests_objective"][i])
  f.write("<|text|>")
  f.write(base_df["quests_description"][i])
  f.write("<|endoftext|>\n")
f.close()
```

Figure 13: Primary code block for pre-processing code

The .txt file was then exported and readied to be loaded into the model. A sample of quests in the new format can be seen in Figure 14.

```
<|startoftext|>Ajunta's Ancient Army<|obj|>Defeat Ancient Droids<|text|>On your way out of the Tomb of Ajunta Pall his ancient droids
activated and attacked. Defeat Ancient Droids in the Tomb of Ajunta Pall<|endoftext|>
<|startoftext|>Arm Yourself<|obj|>Explore the Armory<|text|>Overseer Tremel has brought you to the Sith Academy on Korriban to begin your
trials ahead of schedule. Your first mission is to acquire a blade more appropriate to your destiny-an ancient Sith warblade from the
armory of the tomb of Ajunta Pall.<|endoftext|>
<|startoftext|>Attack of the Flesh Raiders<|obj|>Speak to Master Satele Shan<|text|>Flesh Raiders have mounted an assault on the Jedi
training grounds in the Gnarls. Derrin Weller has asked you to help ward off this threat. Speak to the speeder droid outside of the
Masters' Retreat to find transportation to the Gnarls.<|endoftext|>
<|startoftext|>Claim Your New Identity<|obj|>Speak to Jheeg<|text|>Your first mission as an Imperial agent brings you to Hutta. Your
handler, Keeper, has ordered you to assume a cover identity and convince Nem'ro the Hutt to ally with the Empire. The first step is to
contact Jheeg, an Imperial asset in the town of Jiguuna; he will brief you on your cover. Meet Jheeg at his safe house in Jiguuna.
<|endoftext|>
<|startoftext|>Cleanse the Valley<|obj|>Kill K'lor'slug Foragers<|text|>You have been sent to the tomb of Ajunta Pall as part of your
training at the Sith Academy. Before entering the tomb, you should hone your combat skills on the k'lor'slug foragers in the Valley of the
Dark Lords.<|endoftext|>
```

Figure 14: Sample of SWTOR dataset

Model Approach and Implementation

With the dataset ready to be loaded into the GPT-2 model, we decided to create 2 models to generate text.

Model 1 was created first using the same parameters as the WOW model, but with two notable differences. The first difference was of course to change the training dataset to the SWTOR dataset. Secondly, we changed the restore_from parameter to latest, as we would be fine-tuning the checkpoint we created with the WOW quest dataset. This would mean our total dataset would still include the 24,981 WOW quests in addition to the 2,651 quests from the SWTOR dataset. The benefit to this is that the larger dataset should make it easier to learn

more data of the same structure, and could results in a more understandable model. The downside is that it could result in lore from WOW leaking into quests that are meant to fit within the world of Star Wars. The code for creating Model 1 can be seen in Figure 15.

```
gpt2.finetune(sess2,
              dataset="/content/drive/MyDrive/swtor_final_dataset.txt",
              model_name='124M',
              steps=1000, #number of steps to run for
              restore_from='latest', #fresh means to train from base gpt, use latest to start from an existing checpoint
              run_name='run2', #subfolder within checkpoint to save the model
              print_every=10,#Number of steps to print training progress
              sample_every=200,#Number of steps to print example output
              save_every=500
              )
```

Figure 15: Model 1 Fine-tuning code

Model 2 was created the same way as the WOW model, but just using the SWTOR dataset. As we trained it directly off of the GPT-2 base model, the restore_from parameter was set to fresh, and the total dataset only consists of 2,651 quests, just from Star Wars. We expected the benefit of this to be that it would focus more on the lore of Star Wars, rather than using anything from World of Warcraft. The downside is that the overall dataset ends up being much smaller, and might not be enough for the model to do anything useful with it. The code for Model 2 can be seen in Figure 16.

```
gpt2.finetune(sess3,
              dataset="/content/drive/MyDrive/swtor_final_dataset.txt",
              model_name='124M',
              steps=1000, #number of steps to run for
              restore_from='fresh', #fresh means to train from base gpt, use latest to start from an existing checpoint
              run_name='run3', #subfolder within checkpoint to save the model
              print_every=10,#Number of steps to print training progress
              sample_every=200,#Number of steps to print example output
              save_every=500
              )
```

Figure 16: Model 2 Fine-tuning code

After these fine-tuned checkpoints were created, they were saved and could be loaded anytime the same way the WOW model was.

Results

With the checkpoints created, the model could then be tested. First, some random quests were generated from the models. Model 1's output can be seen below, and was generated as seen in Figure 17.

```
gpt2.generate(sess2, run_name='run2')
```

Figure 17: Random quest generation results for Model 1

>The Shadow Syndicate<|obj|>Find the Second Hostile Situation<|text|>The ICC has become involved in the government of General Aves. Suspicion of General Aves' whereabouts has been tested at this point, but resistance has grown in numbers as the conflict between the Imperials and the Shadow Syndicate grows. General Aves has asked you to find a defector from the General and recover the ICC datapad. A bodyguard is provided as proof of his demise. Retrieve the ICC datapad from the body of a defector.<|endoftext|>
<|startoftext|>The Master's Secret<|obj|>Use the Holoterminal<|text|>At last our targets have been defeated, and the final confrontation between Darth Baras and the Jedi Masters on Taris has begun. The Jedi have traced a crashed dreadnaught to the Voss, and the Empire's advance is dangerously close. The Master of the Voss, Tulak Hord, has asked you to rescue the master's lightsaber. Rance the White Seraph, and obtain the master's holocall. The White Seraph must be in the area. Enter the Voss base.<|endoftext|>
<|startoftext|>The Pilgrim's Path<|obj|>Travel to Taris<|text|>You and Doctor Ianna Cel delivered a holocamera to the Grancha Lakand venom mine, but the location is not known yet. You've been asked to plan a course for a ship belonging to the Chiss hero. Use the holocall generator aboard the ship to draw the attention of the Voss pilgrim Doctora. Use the holocall generator aboard the ship to draw the attention of the Voss pilgrim Doctora.<|endoftext|>
<|startoftext|>A Mystic's Path<|obj|>Speak to Suva-Rak<|text|>You are on the trail of the lost artifacts of Tulak Hord, which are believed to be key to a secret ritual that will grant you great power. Darth Zash has traced one of these artifacts to the Voss. You've arrived on Taris to help a Chiss hero named Suva-Rak fight the Empire. Meet Suva-Rak on Taris.<|endoftext|>
<|startoftext|>A Woman Saved<|obj|>Find the Lost Chiss Chameleon<|text|>Chiss thief Lelo Aure has procured the chiss's valuable toothpail. With it, you can recover time-consuming repairs to old Taris's power grids. Lelo has asked you to steal the toothpail. Locate the toothpail on the other side of the Jundland Compound.<|endoftext|>
<|startoftext|>Away from Home<|obj|>Board the Tularan Spire<|text|>Enter the Imperial base on Nar Shaddaa.<|endoftext|>
<|startoftext|>A Threat from the Past<|obj|>Take a Taxi to the New World<|text|>Your mission on Nar Shaddaa is to create a new, though already mysterious, species of Yuonis. The first, most urgent task is to recover the pieces of a lost world. You've assembled a motley crew of galactic spies, bounty hunters and smugglers to help you. Head to the Drift Hills research camp andart your way to the alien race.<|endoftext|>
<|startoftext|>A Threat from the Past<|obj|>Eliminate the Rakghouls<|text|>Your closest contact, Watcher X, has revealed that a group of ancient rakghouls have invaded the Tularan Spire. This ancient race of aliens may be the most dangerous the Republic has to offer. Take out the rakghouls inside the Tularan Spire on Nar Shaddaa.<|endoftext|>
<|startoftext|>A Threat from the Past<|obj|>Eliminate Republic Saboteurs<|text|>Your closest contact, Watcher X, has revealed that a group of ancient rakghouls have invaded the Tularan Spire. This ancient race may be the most dangerous the Republic has to offer. Take out the Republic saboteurs inside the Tularan Spire on Nar Shaddaa.<|endoftext|>
<|startoftext|>Advanced Training: Agent Galen<|obj|>Report to Agent Agent Nie'al<|text|>You have been sent to the tomb of Ajunta Pall as part of your training at the

Model 2's output can be seen below, and was generated as seen in Figure 18.

```
gpt2.generate(sess3, run_name='run3')
```

Figure 18: Random quest generation results for Model 2

Seek Your Contact<|obj|>Travel to Nar Shaddaa<|text|>You have been contacted by a Data Corp representative who has information that can help you defeat Darth Jadus. The representative, in order to gain access to the data center, has asked you to steal a data spike the representative is using to trap Darth Jadus. Destroy the data spike in the Data Corp outside the Data Corp Mobilization Center.<|endoftext|>
<|startoftext|>Crystal Ball<|obj|>Infiltrate the First Meeting<|text|>A representative from the First Disciple, a powerful warrior cult that stands between the forces of Darth Jadus and the Jedi of old, has decided to sneak a meeting with the leaders of the first to come. The representative, in a bid to hide his identity, has enlisted your help. Place the acoustic disruptors around the arena to drown out the First Disciple.<|endoftext|>
<|startoftext|>First to Disciples<|obj|>Infiltrate the Temple of Corruption<|text|>The meeting with the Jedi will continue in the Fallenspire Stronghold inside the Forge. Diurnal reports indicate that the temblores at the Forge have decided to bring in Jedi stealth troops to protect their generators. You mustn't let that happen. Head to the Fallenspire Stronghold inside the Forge and disable the Jedi's stealth droids to ensure that the generators cannot stand the heat.<|endoftext|>
<|startoftext|>Friends of the Republic<|obj|>Speak with Senator Evran<|text|>A Senator Evran has an important mission for you on Nar Shaddaa. Meet with you bureaus in the Senate tower.<|endoftext|>
<|startoftext|>Bug Hunt<|obj|>Speak with Gearbox Inc.<|text|>You met Dor Ragnan, the head of gaming operations at DICE. He suggested you go into the gaming district and get into some serious tights. He's provided you with the codes to the security terminals to unlock the secured consoles. Go into the gaming district and get into some serious tights.<|endoftext|>
<|startoftext|>Bugged-Out<|obj|>Revert Warden Droids<|text|>You've been instructed to reverse engineer the security droids in the warden posts to make sure the bugs haven't caused any injuries to anyone else. Turn in the Warden Droids data spike at the drop box in the Maximum Security Section.<|endoftext|>
<|startoftext|>Bust Up Skelton<|obj|>Defeat Warden Slaughter<|text|>Warden Slaughter<|endoftext|>
<|startoftext|>Clean Up the Cell Block<|obj|>Kill K'lor'slug Forces<|text|>While activating the baradium detonator in the Fallenspire Stronghold, you were attacked by k'lor'slugs. Defeat the k'lor'slug aggression as you locate the chamber leading to the chamber of the k'lor'slug's foul-mouthed benefactor.<|endoftext|>
<|startoftext|>Clean Up the TRANPCOT<|obj|>Defeat Killik Forces<|text|>As you destroy the k'lor'slug tunnels, take out any Killiks that stand in your way.<|endoftext|>
<|startoftext|>Cleaning up the Mess<|obj|>Plant a Mouse Droid on the Killik Hive<|text|>You have been tasked with wiping out the Killiks that have infested the area. As you hunt down the bugs, place a mouse droid on the Killik Hive tank to bring their numbers down a notch. Plant a mouse droid on the Killik Hive tank to bring its defense down a notch. Next, defeat Killik forces inside the Ruur Killik Burrows.<|endoftext|>
<|startoftext|>Clear Out the Pipes<|obj|>Kill Killik Forces<|text|>You have been tasked with wiping out the Killiks that have infested the area. As you hunt down the bugs, place a mouse droid on the Killik Hive tank to bring their numbers down a notch. Eliminate Killik forces inside the Ruur Killik Burrows.<|endoftext|>
<|startoftext|>Clearing the Council Chamber<|obj|>Defeat Imperial Forces<|text|>Imperial forces have fortified their position inside the Jedi Temple. Gain extra access to the Jedi

Council by defeating

Next, we generate text based off of a specific quest title and objective. In this example, we used the quest title ‘The Frozen War’ and the quest objective ‘Talk to the wampas on Hoth.’. This title and objective was expected to generate quest descriptions that would fit within the context of talking to wampas on the planet Hoth (in the lore of Star Wars, wampas are animalistic creatures that reside on the snowy planet of Hoth). The code to generate this from Model 1 can be seen in Figure 19, and its outputs can be seen below.

```
gpt2.generate(sess2,
              length=250, #number of tokens to generate
              temperature=0.7,#level of text unpredictability
              prefix="<|startoftext|>The Frozen War<|obj|>Talk to the wampas on Hoth.<|text|>", #prefix to generate text off of
              truncate="<|endoftext|>", #truncate input text when encountering this sequence
              nsamples=5, #number of output text samples to show
              batch_size=5,
              include_prefix=False, #hides prefix in output sample text
              )
```

Figure 19: Generating text for the title-objective prompt in Model 1

Your mission on Hoth is to recruit a technical specialist named Yuun into Havoc Squad. Yuun is currently assigned to recreate an Umbra encrypter, the device responsible for decades of unbreakable Imperial codes, using parts from crashed warships spread across Hoth’s frozen wastes. Sergeant Yuun has determined that the final Umbra component has been taken to the Star of Coruscant, a fallen dreadnaught occupied by the White Maw pirates. Unfortunately, Sergeant Yuun’s squad has suffered heavy casualties, and they are unable to assist you in recovering the component. Travel to the Star of Coruscant in the Starship Graveyard and recover the final Umbra component.
==========
The Republic is fighting to win favor with the Voss, the inhabitants of a newly discovered and mysterious world. You have been asked to assist Ambassador Jannik in his efforts to negotiate an alliance. After exposing a Sith plot to recruit failed Mystics, the Voss envoy Laryn-Ki revealed that the Mystics have had a vision. You will help the Voss defeat the Gormak, a race of technologically advanced warriors and the mortal enemies of the Voss people.
==========
You have been sent by Darth Jadus to rescue the Voice of the Emperor. You have defeated the Flesh Raiders and learned they have a nearby ruin. Go to the Voice of the Emperor rally at Jadus’s camp in the Glacial Fissure.
==========
You have been sent by the Emperor’s Hand to rescue the Voice of the Emperor. The reclusive wampas on Hoth are assembling to mark one of the crystal pillars as a symbol of Imperial might. The best place to start is the prepared warren inside the Jedi Temple. Infiltrate the Three Wispers Jedi Temple, and talk to the wampas there.
==========
You have come to Hoth to rescue the Voice of the Emperor. Madjar Thul’s entourage includes Lieutenant Poh, a former Imperial soldier who should know no more. Speak to Lieutenant Poh at his villa in the Gorinth Canyon to find out what’s going on.
==========

The code to generate this from Model 2 can be seen in Figure 20, and its outputs can be seen below.

```
gpt2.generate(sess3,
              run_name='run3',
              length=250, #number of tokens to generate
              temperature=0.7,#level of text unpredictability
              prefix="<|startoftext|>The Frozen War<|obj|>Talk to the wampas on Hoth.<|text|>", #prefix to generate text off of
              truncate="<|endoftext|>", #truncate input text when encountering this sequence
              nsamples=5, #number of output text samples to show
              batch_size=5,
              include_prefix=False, #hides prefix in output sample text
              )
```

Figure 20: Generating text for the title-objective prompt in Model 1

Evaluation of Models

In terms of evaluating the performance of the model, the original paper used surveys from anonymous individuals to evaluate the quality of the outputs of the model (van Stegeren & Myśliwiec, 2021). This was done by creating an online survey and asking participants to evaluate a list of 20 quests with their titles, objectives and descriptions (van Stegeren & Myśliwiec, 2021). 10 of these quests were collected from the WOW database and 10 were generated by the model. They were evaluated on the merits of English language quality, coherence in context of the quest, novelty of the idea, creativity and level of surprise (van Stegeren & Myśliwiec, 2021). Due to time constraints, it was not feasible for us to do a similar survey to determine how others might perceive our results and measure the similarity of results. However, if we were to do a similar survey, we would create a similar survey to the original paper, but we would include our outputs for the quest title “In Dire Need” with those of the original paper’s. We would then compare the overall metrics on the same metrics used in the original paper’s survey to observe how our model compared to the original. We would then follow up with a second survey to work on our Star Wars models. In this case, we would generate a variety of different quests with different titles, descriptions and objectives, and would provide a survey on 15 different quests. This would consist of 3 sets: 5 quests generated by Model 1, 5 quests generated by Model 2 and 5 quests from the original database. The user would not be told which quests are from which set, but would be asked to rate each quest using the same metrics and questioning as for the WOW dataset.

Conclusion

This project provided us with useful experience for fine-tuning an existing GPT model on new datasets. After reproducing the work done in the existing paper, we created a new dataset based on the SWTOR dataset, and fine-tuned two new models: one based on the existing checkpoint created in the reproduction stage, and a second model fine-tuned directly on the GPT base model.

If we were to perform further work on this project, we would first want to perform the evaluation methods discussed previously. It would be important to determine how our models compare with the existing methods and datasets before going further with it. Despite the

subjective metrics like asking someone to evaluate the results, objective metrics such as perplexity, BLEU score, and other language modeling evaluation metrics can be used to measure the accuracy and coherence of the generated text. Following this, we would also attempt to use a larger dataset more comparable to the WOW database size. In order to do this, we would also need to determine a new solution for the web parsing as the free version of ParseHub only allows parsing of 200 links at a time, which was quite time-consuming and inefficient for this work. Further down the line, it could also be interesting to add other tags to quests that could be parsed, such as the in-game location or recommended player level. We would also want to explore further options for varying the fine-tuning parameters to see what the results may look like.

There is much that could be done to add to this project in the future, but the work we have done so far was quite interesting and insightful for our team.

References

Ashby, T., Webb, B., Knapp, G., Searle, J., & Fulda, N. (2023). Personalized Quest and Dialogue Generation in Role-Playing Games: A Knowledge Graph-and Language Model-based Approach. 23. https://doi.org/10.1145/3544548.3581441

Better language models and their implications. (n.d.). Openai.com. https://openai.com/research/better-language-models

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., & Hesse, C. (2020). Language Models are Few-Shot Learners. Arxiv.org. https://arxiv.org/abs/2005.14165

Bucinca, Z., Yemez, Y., Erzin, E., & Sezgin, M. (2020). AffectON: Incorporating Affect Into Dialog Generation. ArXiv:2012.06847 [Cs]. https://arxiv.org/abs/2012.06847

Dai, K. (2020). Multi-Context Dependent Natural Text Generation for More Robust NPC Dialogue. Dash.harvard.edu. https://nrs.harvard.edu/URN-3:HUL.INSTREPOS:37364681Hämäläinen, M., Alnajjar, K., & Poibeau, T. (2022). Video Games as a Corpus: Sentiment Analysis using Fallout New Vegas Dialog. FDG '22: Proceedings of the 17th International Conference on the Foundations of Digital Games. https://doi.org/10.1145/3555858.3555930

Kacmarcik, G. (2021). Using Natural Language to Manage NPC Dialog. Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment, 2(1), 115–117. https://doi.org/10.1609/aiide.v2i1.18757

Kalbiyev, A. (2022, January 19). Affective dialogue generation for video games. Essay.utwente.nl. http://essay.utwente.nl/89325/

Kerr, C., & Szafron, D. (2009). Supporting Dialogue Generation for Story-Based Games. Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment, 5(1), 154–160. https://doi.org/10.1609/aiide.v5i1.12371

Mateas, M., & Stern, A. (2003). Façade: An Experiment in Building a Fully-Realized Interactive Drama.

OpenAI. (2023). GPT-4 Technical Report. ArXiv:2303.08774 [Cs]. https://arxiv.org/abs/2303.08774v3

ParseHub. (n.d.). Retrieved April 21, 2023, from https://www.parsehub.com/

Quests. SWTOR Spy. (n.d.a). Retrieved April 19, 2023, from http://www.swtor-spy.com/quests/

Quests. Wowhead. (n.d.b). Retrieved April 6, 2023, from https://www.wowhead.com/quests

Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language Models are Unsupervised Multitask Learners. *OpenAI Blog*, 1(8). https://doi.org/https://d4mucfpksywv.cloudfront.net/better-language-models/language-models.pdf

van Stegeren, J., & Myśliwiec, J. (2021). Fine-tuning GPT-2 on annotated RPG quests for NPC dialogue generation. *The 16th International Conference on the Foundations of Digital Games (FDG) 2021, 2*, 1–8. https://doi.org/https://doi.org/10.1145/3472538.3472595

Woolf, M. (2019, September 4). *How to make custom AI-generated text with GPT-2*. Max Woolf's Blog. Retrieved April 3, 2023, from https://minimaxir.com/2019/09/howto-gpt2/