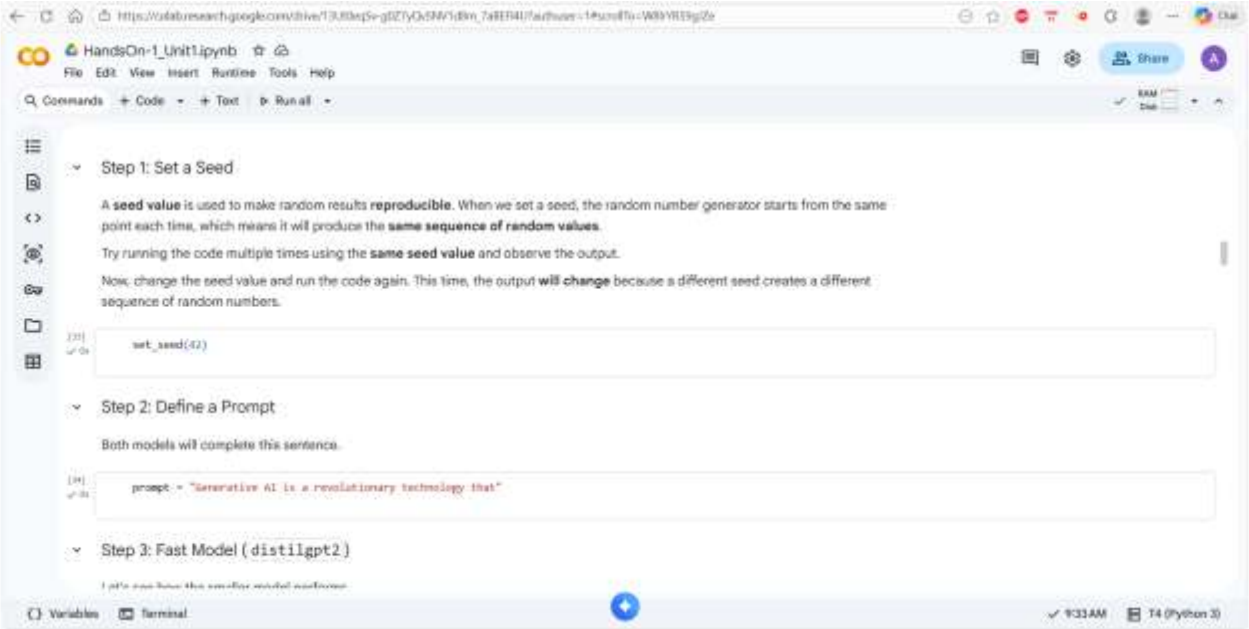
GENAI HANDSON

Name: Ananya Reddy Gandlaparthi

SRN: PES2UG23CS064

Section: A


HandsOn-1_Unit1.ipynb
File Edit View Insert Runtime Tools Help
Commands + Code + Text Run all
Share
Variables Terminal
9:33 AM T4 (Python 3)

Step 1: Set a Seed

A **seed value** is used to make random results **reproducible**. When we set a seed, the random number generator starts from the same point each time, which means it will produce the **same sequence of random values**.

Try running the code multiple times using the **same seed value** and observe the output.

Now, change the seed value and run the code again. This time, the output **will change** because a different seed creates a different sequence of random numbers.

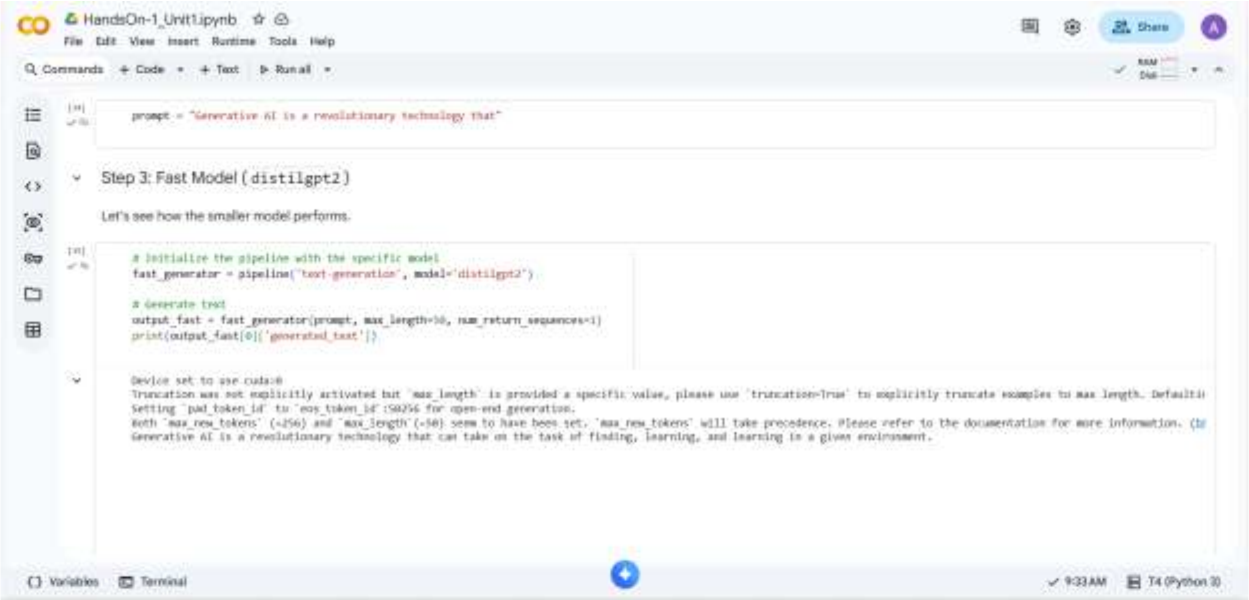
```
set_seed(42)
```

Step 2: Define a Prompt

Both models will complete this sentence.

```
prompt = "Generative AI is a revolutionary technology that"
```

Step 3: Fast Model (distilgpt2)


HandsOn-1_Unit1.ipynb
File Edit View Insert Runtime Tools Help
Commands + Code + Text Run all
Share
Variables Terminal
9:33 AM T4 (Python 3)

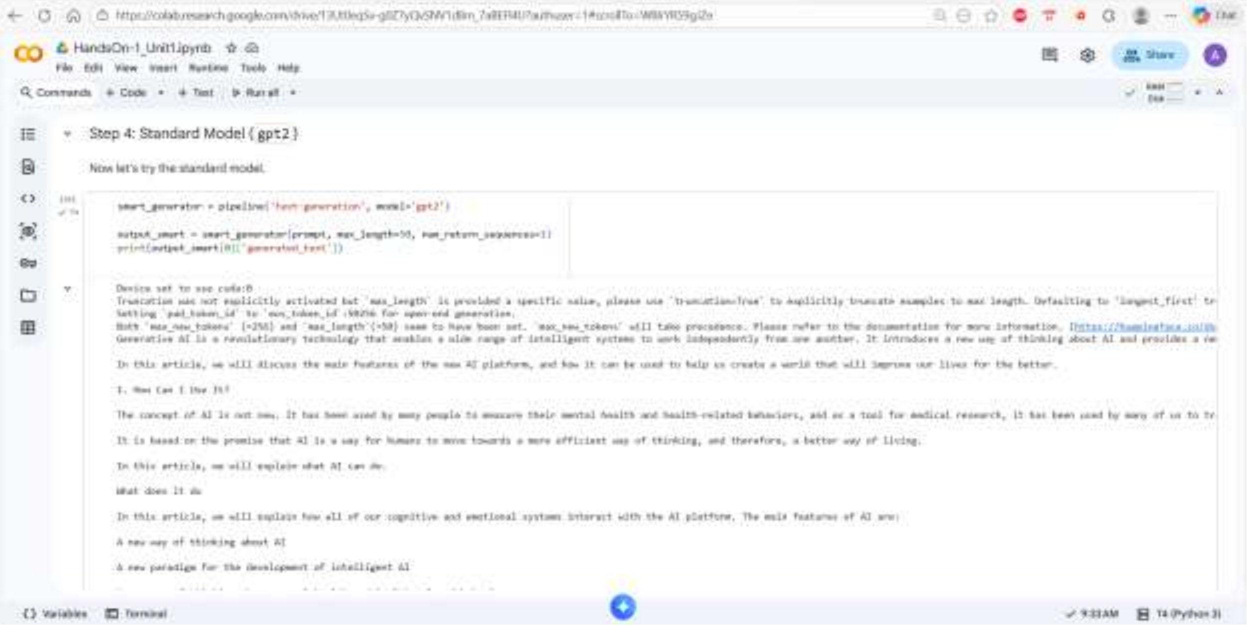
```
prompt = "Generative AI is a revolutionary technology that"
```

Step 3: Fast Model (distilgpt2)

Let's see how the smaller model performs.

```
# Initialize the pipeline with the specific model
fast_generator = pipeline('text-generation', model='distilgpt2')

# Generate text
output_fast = fast_generator(prompt, max_length=50, num_return_sequences=1)
print(output_fast[0]['generated_text'])
```

```
Device set to use cuda:0
Truncation was not explicitly activated but `max_length` is provided a specific value, please use `truncation=True` to explicitly truncate examples to max length. Defaulti
Setting `pad_token_id` to `eos_token_id`:50256 for open-end generation.
Both `max_new_tokens` (=256) and `max_length`(=50) seem to have been set. `max_new_tokens` will take precedence. Please refer to the documentation for more information. (h
Generative AI is a revolutionary technology that can take on the task of finding, learning, and learning in a given environment.
```

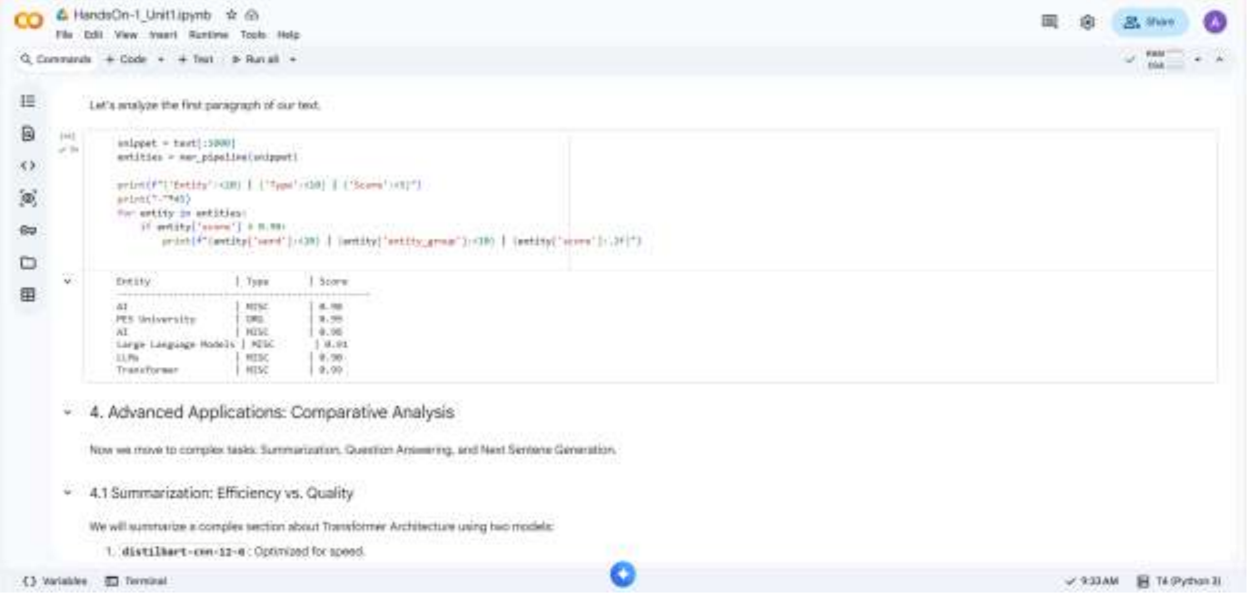

https://colab.research.google.com/drive/1...
HandsOn-1_Unit1.ipynb
File Edit View Insert Runtime Tools Help
Commands + Code + Text ▷ Run all
Share
Variables Terminal
✓ 9:33 AM T4 (Python 3)

Step 4: Standard Model (gpt2)

Now let's try the standard model.

```
smart_generator = pipeline('text-generation', model='gpt2')

output_smart = smart_generator(prompt, max_length=50, num_return_sequences=1)
print(output_smart[0]['generated_text'])
```

```
Device set to use cuda:0
Truncation was not explicitly activated but `max_length` is provided a specific value, please use `truncation=True` to explicitly truncate examples to max length. Defaulting to 'longest_first' tr
Setting `pad_token_id` to `eos_token_id`:50256 for open-end generation.
Both `max_new_tokens` (=256) and `max_length`(=50) seem to have been set. `max_new_tokens` will take precedence. Please refer to the documentation for more information. (https://huggingface.co/do
Generative AI is a revolutionary technology that enables a wide range of intelligent systems to work independently from one another. It introduces a new way of thinking about AI and provides a ne

In this article, we will discuss the main features of the new AI platform, and how it can be used to help us create a world that will improve our lives for the better.

1. How Can I Use It?

The concept of AI is not new. It has been used by many people to measure their mental health and health-related behaviors, and as a tool for medical research, it has been used by many of us to tr

It is based on the premise that AI is a way for humans to move towards a more efficient way of thinking, and therefore, a better way of living.

In this article, we will explain what AI can do.

What does it do

In this article, we will explain how all of our cognitive and emotional systems interact with the AI platform. The main features of AI are:

A new way of thinking about AI

A new paradigm for the development of intelligent AI
```


HandsOn-1_Unit1.ipynb
File Edit View Insert Runtime Tools Help
Commands + Code + Text ▷ Run all
Share
Variables Terminal
✓ 9:33 AM T4 (Python 3)

Let's analyze the first paragraph of our text.

```
snippet = text[:1000]
entities = ner_pipeline(snippet)

print(f"{'Entity':<20} | {'Type':<10} | {'Score':<5}")
print("-"*45)
for entity in entities:
    if entity['score'] > 0.90:
        print(f"{entity['word']:<20} | {entity['entity_group']:<10} | {entity['score']:.2f}")
```

```
Entity               | Type       | Score
---------------------------------------------
AI                   | MISC       | 0.98
PES University       | ORG        | 0.99
AI                   | MISC       | 0.98
Large Language Models | MISC       | 0.91
LLMs                 | MISC       | 0.98
Transformer          | MISC       | 0.99
```

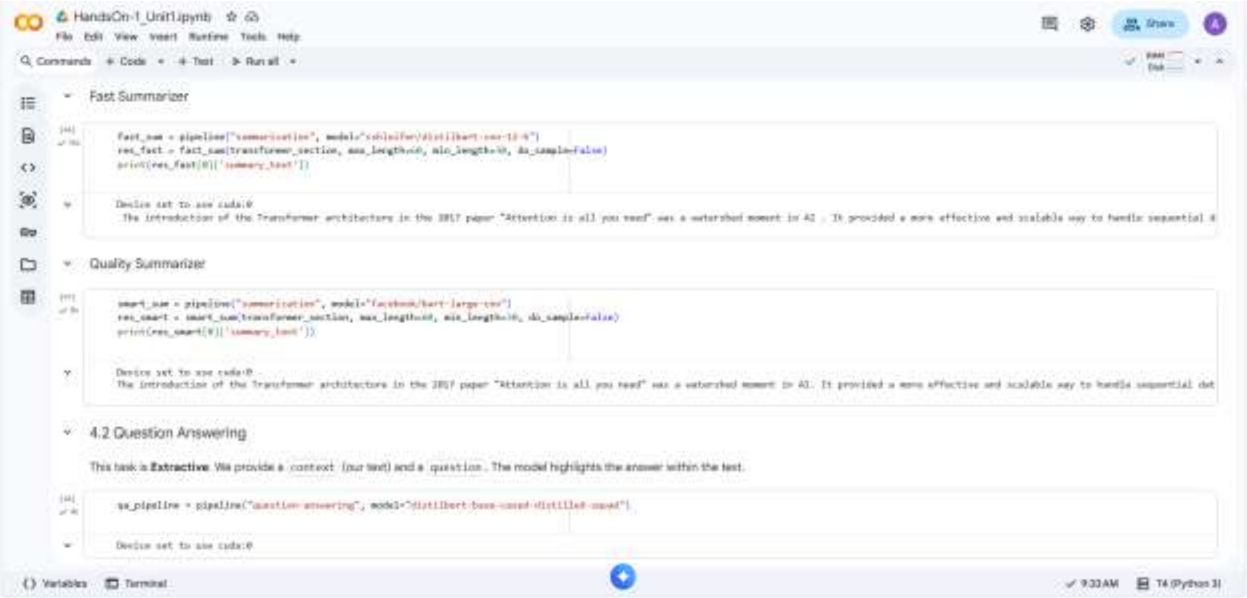
4. Advanced Applications: Comparative Analysis

Now we move to complex tasks: Summarization, Question Answering, and Next Sentene Generation.

4.1 Summarization: Efficiency vs. Quality

We will summarize a complex section about Transformer Architecture using two models:

1. `distilbart-cnn-12-6` : Optimized for speed.


HandsOn-1_Unit1.ipynb
File Edit View Insert Runtime Tools Help
Commands + Code + Text ▶ Run all
Share
Variables
Terminal
9:33 AM
T4 (Python 3)

Fast Summarizer

```
fast_sum = pipeline("summarization", model="sshleifer/distilbart-cnn-12-6")
res_fast = fast_sum(transformer_section, max_length=60, min_length=30, do_sample=False)
print(res_fast[0]['summary_text'])
```

```
Device set to use cuda:0
 The introduction of the Transformer architecture in the 2017 paper "Attention is all you need" was a watershed moment in AI . It provided a more effective and scalable way to handle sequential d
```

Quality Summarizer

```
smart_sum = pipeline("summarization", model="facebook/bart-large-cnn")
res_smart = smart_sum(transformer_section, max_length=60, min_length=30, do_sample=False)
print(res_smart[0]['summary_text'])
```

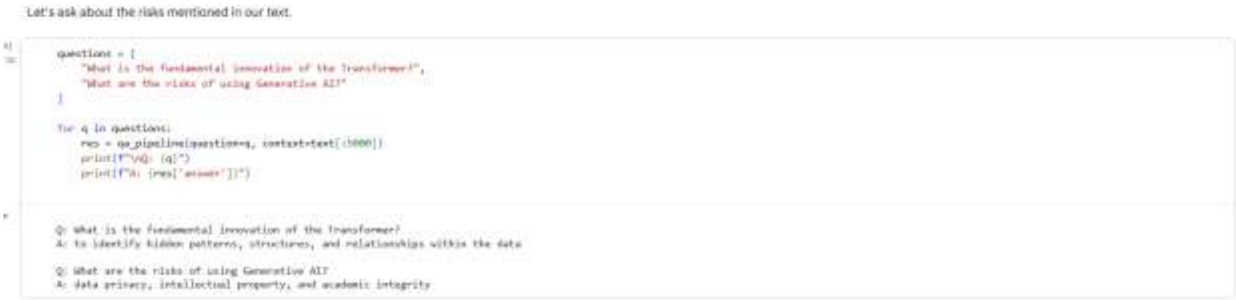
```
Device set to use cuda:0
The introduction of the Transformer architecture in the 2017 paper "Attention is all you need" was a watershed moment in AI. It provided a more effective and scalable way to handle sequential dat
```

4.2 Question Answering

This task is **Extractive**. We provide a `context` (our text) and a `question`. The model highlights the answer within the text.

```
qa_pipeline = pipeline("question-answering", model="distilbert-base-cased-distilled-squad")
```

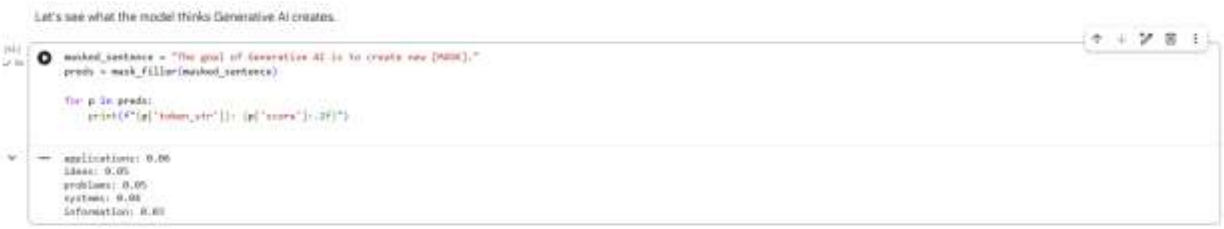
```
Device set to use cuda:0
```



Let's ask about the risks mentioned in our text.

```
questions = [
    "What is the fundamental innovation of the Transformer?",
    "What are the risks of using Generative AI?"
]

for q in questions:
    res = qa_pipeline(question=q, context=text[:5000])
    print(f"\nQ: {q}")
    print(f"A: {res['answer']}")
```

```
Q: What is the fundamental innovation of the Transformer?
A: to identify hidden patterns, structures, and relationships within the data

Q: What are the risks of using Generative AI?
A: data privacy, intellectual property, and academic integrity
```

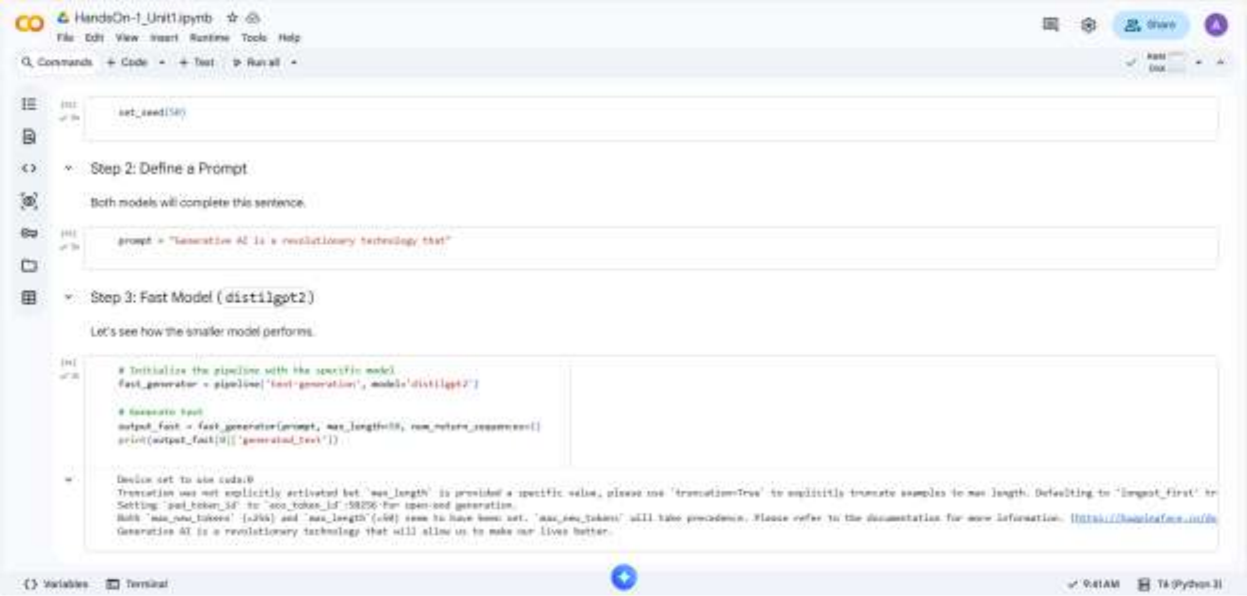


Let's see what the model thinks Generative AI creates.

```
masked_sentence = "The goal of Generative AI is to create new [MASK]."
preds = mask_filler(masked_sentence)

for p in preds:
    print(f"{p['token_str']}: {p['score']:.2f}")
```

```
applications: 0.06
ideas: 0.05
problems: 0.05
systems: 0.04
information: 0.03
```

Seed: 50



HandsOn-1_Unit1.ipynb
File Edit View Insert Runtime Tools Help
Commands + Code + Text Run all
Share

```
set_seed(50)
```

Step 2: Define a Prompt

Both models will complete this sentence.

```
prompt = "Generative AI is a revolutionary technology that"
```

Step 3: Fast Model (distilgpt2)

Let's see how the smaller model performs.

```
# Initialize the pipeline with the specific model
fast_generator = pipeline('text-generation', model='distilgpt2')

# Generate text
output_fast = fast_generator(prompt, max_length=50, num_return_sequences=1)
print(output_fast[0]['generated_text'])
```

```
Device set to use cuda:0
Truncation was not explicitly activated but `max_length` is provided a specific value, please use `truncation=True` to explicitly truncate examples to max length. Defaulting to 'longest_first' tr
Setting `pad_token_id` to `eos_token_id`:50256 for open-end generation.
Both `max_new_tokens` (=256) and `max_length`(=50) seem to have been set. `max_new_tokens` will take precedence. Please refer to the documentation for more information. (https://huggingface.co/do
Generative AI is a revolutionary technology that will allow us to make our lives better.
```

The distilgpt2 model, which is smaller and faster, generated a short, simple sentence. It was quick and gave a basic answer, but the output was limited. GPT2 model, which is bigger and smarter, generated much longer text with more details about AI benefits. It stayed more on topic and provided more information. The seed value of 42 made sure that both runs produced the same output each time.

Using the same seed value of 42 gave me the same generated text every time I ran the code, but when I changed the seed to 50 I got a different output. Bigger models like gpt2 and bart-large produce much better quality outputs than smaller models like distilgpt2 and distilbart, but they are slower and require more computing power.