Reflection Report - Exploring GPT-2 for Text Generation using Transformers

1. Introduction

Transformers are a type of deep learning model designed to handle sequential data. Introduced in the paper *"Attention is All You Need"*, the Transformer architecture uses self-attention to weigh the importance of each word in a sequence, regardless of its position.

Transformers are crucial in modern AI because they enable powerful language understanding and generation. Models like GPT-2 can generate realistic and coherent text, making them useful for tasks such as chatbots, content creation, translation, and summarization.

2. Experiment Summary

In this project, I used the **GPT-2** model from Hugging Face's transformers library within a Jupyter Notebook in **VS Code**. I generated text based on various prompts using the pipeline("text-generation") API and experimented with different hyperparameters like:

- temperature (controls randomness)
- max_length and max_new_tokens
- repetition_penalty
- top_k and top_p
- num_return_sequences

Prompts I Tried:

- Futuristic statement: "In the future, education will"
- News headline: "Breaking news: Artificial intelligence has surpassed..."
- Story starter: "Once upon a time, in a world ruled by robots..."
- Dialogue: "Person A: What do you think about..."
- Question: "What are the implications of quantum computing..."
- Open phrase with multiple outputs: "The future of transportation includes"

---

3. Observations

💠 Prompt 1: "In the future, education will"

- **Output:** GPT-2 generated a paragraph about democratizing education, tuition struggles, and government policy.
- **Observation:** The text mimicked political discourse, with realistic names and references. It occasionally drifted into hallucinated or contradictory facts.

💠 Prompt 2: "Breaking news..."

- **Output:** GPT-2 generated a news-style paragraph referencing AI in medical diagnostics and robotic prosthetics.
- **Observation:** The model mimicked journalism well, even creating ads and quotes. However, some factual accuracy was lacking or nonsensical.

💠 Prompt 3: "Once upon a time..."

- **Output:** The model generated a story involving vampires, cults, and strange physical descriptions.
- **Observation:** This was an unexpectedly dark and surreal result. It shows GPT-2's tendency to hallucinate vivid but incoherent fiction, especially with higher temperature and storytelling prompts.

💠 Prompt 4: Dialogue

- **Output:** A back-and-forth conversation with rich vocabulary, discussing power, corporations, and revolutions.
- **Observation:** The output was surprisingly deep and philosophical, but also confusing. Dialogue generation shows GPT-2 can simulate multiple speakers, though often disjointedly.

💠 Prompt 5: Quantum Computing

- **Output:** A repetitive block stating "how the universe works" dozens of times.

- **Observation:** Repetition was a major issue. This likely occurred due to a low temperature and no repetition penalty. GPT-2 sometimes loops without new content.

💠 **Prompt 6: Transportation (3 outputs)**

- **Output:** GPT-2 generated varied outputs ranging from hybrid vehicles to emission standards and rail expansion.
- **Observation:** This was the most consistent type of output. The language was formal and policy-oriented, akin to real-world infrastructure reports.

---

4. Reflection

This project provided hands-on insight into how GPT-2 models generate text based on patterns, associations, and training data frequency. I observed that:

- **Changing temperature** altered randomness: lower values led to boring/repetitive content, while higher values created more diverse but sometimes incoherent results.
- **Prompt phrasing** deeply affected tone and structure: even slight word changes influenced the theme.
- **Repetition** was a common limitation, especially for questions and open-ended prompts.
- GPT-2 lacks true "understanding", it doesn't reason or track context long-term.

Surprising Findings:

- The storytelling output featured dark fictional elements with unexpected creativity.
- GPT-2 produced very realistic-sounding bureaucratic or political language with minimal prompt hints.

Limitations:

- Repetition and hallucination.
- Lack of deep reasoning.
- Doesn't maintain logical flow in longer outputs or complex narratives.

5. Sample Outputs

Prompt 1: *In the future, education will*
Result:

> "...be a key element in the development of our democratic system. We need to build a democratic university... the government is looking at ways to help students with disabilities..."

Prompt 2: *Breaking news: Artificial intelligence has surpassed human intelligence...*
Result:

> "...Experts say it's about to overtake everything else. A video showing a robotic arm being used..."

Prompt 3: *Once upon a time, in a world ruled by robots...*
Result:

> "...a young girl discovered she had lost her virginity after she had been tricked into being a vampire... she joined a cult..."

Prompt 4: *What are the implications of quantum computing...*
Result:

> "...we have to understand how the universe works to understand how it works to understand how it works..." (repeats endlessly)

Prompt 5: *The future of transportation includes*
Result:

> "Efficient hybrid drive systems, such as electric vehicles... U.S. Environmental Protection Agency (EPA)... national rail network..."