

Sentiment Analysis of Image Generation Prompts

Stav Nimrod

June 2023

1 Abstract

This study investigates the sentiment of image generation prompts and explores patterns related to sentiment in the prompts' properties and the generated images. The research process includes cleaning and preprocessing text prompts, performing sentiment analysis, analyzing sentiment score distributions, comparing prompt lengths, examining word frequencies, creating word clouds, and extracting color information from generated images. The findings reveal intriguing relationships between prompt properties and sentiment scores, providing valuable insights for future studies in image generation and natural language processing.

2 Introduction

In recent years, image generation models have made significant advancements, with text prompts playing a crucial role in guiding the generation process. These prompts, often in the form of sentences or short phrases, determine the content, style, and, in many cases, the sentiment of the resulting images. Understanding the characteristics of these prompts and their impact on the generated images is vital for advancing the field of image generation. This paper focuses on sentiment analysis of image generation prompts, aiming to gain insights into the prompt properties and their influence on the output images. By exploring the relationship between sentiment and prompt characteristics, we can shed light on how sentiment is reflected in the generated images and pave the way for further research in image generation and natural language processing.



Figure 1: Generated Image using Midjourney (Prompt: a realistic cow that is flying in the sky with goggles, small wings, clouds)

3 Related Work

Several researchers have focused on sentiment analysis in different domains like product reviews, social media, and more. However, the exploration of sentiment analysis in image generation prompts is still a relatively untouched area.

4 Methodology

The dataset used in this study consists of image generation prompts obtained from the Midjourney text-to-image service dataset (Midjourney User Prompts and Generated Images) [1]. Midjourney, an independent research lab, introduced a text-to-image service in 2022 [2], allowing users to interact with a Midjourney bot on a public Discord server to issue queries and receive corresponding images. The dataset was collected by scraping messages from the public Discord server over a period of four weeks (June 20, 2022 - July 17, 2022).

The dataset initially contained 250,000 generated images. However, upon closer examination, it was observed that only 145,080 images were generated from the original prompts, while the remaining images were generated as a result of user requests for image upscaling or creating variants of already generated images. In order to maintain consistency and focus on the analysis of initial prompts and images, only the data related to these initial prompts and images was considered for this study.

Upon analyzing the data, it became evident that there were numerous instances of prompts that exhibited slight variations. It appeared that users became more proficient in formulating prompts, incorporating key words, referencing specific images, and incorporating unique parameters to control the style, aspect ratio, and other aspects of the generated images. Despite the presence of these slightly altered prompts, it was observed that the overall analysis conclusions remained consistent. Consequently, it was decided to retain these prompts in the dataset rather than removing them, as they contributed valuable insights to the analysis.

To conduct sentiment analysis, the Hugging Face Transformers library with the sentiment-analysis pipeline [3] was utilized. This library employs a pre-trained model fine-tuned for sentiment classification and provides efficient GPU processing capabilities. The image generation prompts were subjected to pre-processing, including the removal of URLs, prompt parameters, and specific patterns. The sentiment analysis results were then appended to the dataset, enabling the analysis of sentiment score distributions.

Additionally, the text prompts underwent tokenization and normalization, with the removal of stop words. Word frequencies were computed for both positive and negative sentiment categories, as well as for the entire dataset. By comparing the differences in word frequencies between the positive and negative categories, the most characteristic words for each sentiment category were identified.

Furthermore, the images generated from the prompts were analyzed to identify their dominant color features. This analysis involved employing K-Means clustering to determine the top colors associated with each sentiment category. By comparing the dominant colors, further insights into the relationship between sentiment and color in the generated images were obtained.

Overall, the methodology employed in this study ensured comprehensive analysis of the image generation prompts, incorporating sentiment analysis, word frequency examination, and color analysis. The careful selection and preprocessing of the dataset allowed for meaningful and consistent insights to be derived from the data.

5 Results and Discussion

5.1 Distribution of Sentiment Scores

The analysis of sentiment scores revealed a bimodal distribution, with distinct peaks in two separate ranges. Negative sentiment scores were primarily concentrated in the range of -1 to -0.9, continuing to -0.5, while positive sentiment scores were predominantly observed in the range of 1 to 0.9, continuing to 0.5 (see Figure 2).

This bimodal distribution of sentiment scores carries significant implications. It suggests the existence of two distinct sentiment categories within the image generation prompts. The separation of sentiment scores into two clusters indicates the presence of contrasting sentiment polarities: negative and positive. The distinct peaks in each range imply a significant preference for sentiments on either end of the spectrum, with fewer prompts expressing sentiments closer to neutral.

Understanding the bimodal distribution of sentiment scores provides insights into the sentiment dynamics within image generation prompts. It highlights the pronounced influence of sentiment polarity in guiding the generation process and the significance of eliciting and capturing specific emotional responses through prompt design. This finding suggests the importance of aligning the sentiment objectives and desired emotional impact with the corresponding prompt properties in image generation tasks.

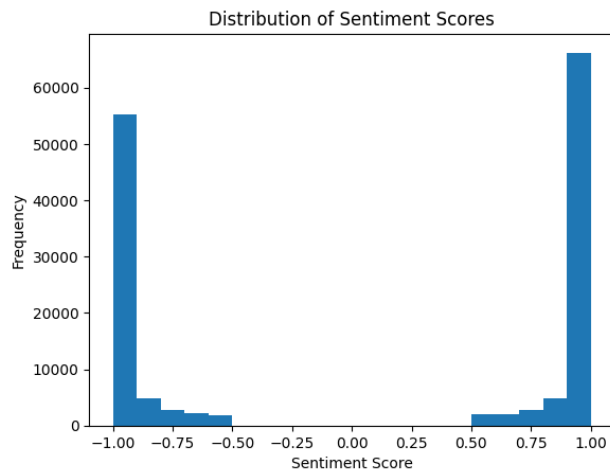


Figure 2: Distribution of Sentiment Scores

5.2 Prompt Lengths and Sentiment

The analysis of prompt lengths provided interesting insights into the relationship between sentiment and prompt characteristics. It was observed that prompts with positive sentiment tended to be longer in length compared to those with negative sentiment. This finding suggests that individuals expressing positive sentiments in their image generation prompts may be inclined to provide more detailed and elaborate descriptions. The additional length allows for a more comprehensive and nuanced expression of their desired visual concepts. On the other hand, prompts with negative sentiment, which were typically shorter, might indicate a more concise

and straightforward expression of the desired image. These findings emphasize the role of prompt length as a potential indicator of sentiment and highlight the differences in the level of detail and complexity between positive and negative image generation prompts.

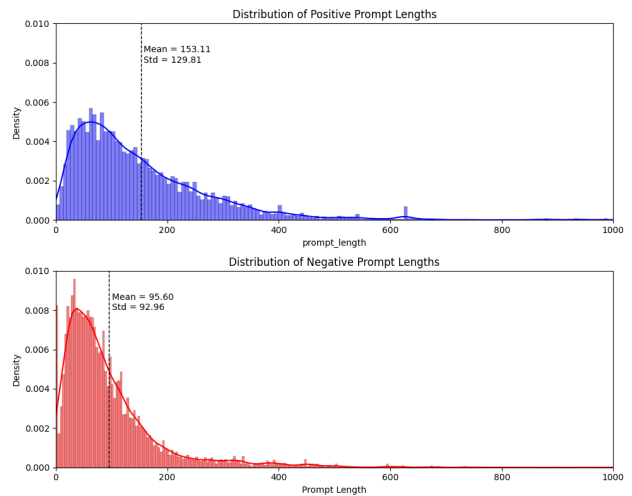


Figure 3: Distribution of Prompt Lengths

5.3 Word Frequencies and Sentiment

Word frequency analysis of the prompt data revealed distinctive words associated with positive and negative sentiments. In positive prompts, words such as "Beautiful," "Cinematic," and "Symmetrical" were frequently observed, indicating a preference for visually appealing and harmonious elements. On the other hand, negative prompts exhibited a higher frequency of words like "Old," "Black," and "Metal," suggesting a tendency towards darker or aged visual concepts. These findings highlight the influence of specific words on sentiment and provide insights into the preferences and characteristics of positive and negative image generation prompts.

7 Acknowledgement

The author would like to express gratitude to Dr. Michael Fire for guiding this research as part of Mining Massive Datasets Course at Ben-Gurion University of the Negev.

A special note of thanks is also extended to those who indirectly contributed to this study. Their active participation in the public Discord server and creation of a diverse range of image generation prompts has greatly enriched this research.

8 References

1. **Midjourney User Prompts and Generated Images (250k).** Retrieved from:
<https://www.kaggle.com/datasets/succinctlyai/midjourney-texttoimage>
2. **Midjourney.** Retrieved from:
<https://www.midjourney.com>
3. **Hugging Face Transformers library with the sentiment-analysis pipeline.** Retrieved from:
https://huggingface.co/docs/transformers/main_classes/pipelines