Towards Better Connection with Food Insecure Communities: Analyzing Language and Tone of Text Generated by AI Tools

Maya Papayan[1](Author), Dr. Benito Mendoza[1], Dr. Marcelo Sztainberg[2], Dr. Nancy Wrinkle[2](Advisors)

New York City College of Technology, New York, NY 10017[1]

Northeastern Illinois University, Chicago, IL 60625[2]

Abstract

According to the USDA, in 2021, about 38 million people in the US were food insecure, with the Hispanic population leading the charts. For example, one in every four people who suffer food insecurity is Hispanic in New York City. In Chicago, this number is even higher, one in every three people. Food organizations use different information systems, such as social networks, websites, and apps, to reach diverse communities. However, some communities are still disconnected, and food insecurity remains high. One of the reasons for this problem is the emotional disconnection between the communities and the language used by these organizations. The work presented here is part of a larger project that aims to understand which language emotions connect better with the communities. Here, we compare AI-based text generation tools such as OpenAI, Rytr.me, Writesonic, and CopyAI. Using these tools, which include options for selecting different emotional tones, such as "exciting," "compassionate," "convincing," etc., we generate and analyze text promoting food services. Using the "Bag of Words" approach, we analyze the similarities among the lexicons utilized by each tool. These results can help us to determine which text generator is well-suited for our cause and help food organizations to improve the engagement rate with their audience.

Introduction

According to USDA 10.2% of U.S. households (13.5 million) were food insecure at some time during 2021. One in every six (17%) food insecure households across the U.S. were Hispanic. In addition, in the areas with the higher Hispanic population density, these numbers were much higher. For example, in New York City, 29% (roughly 1 in 3) of food insecure households were reported Hispanic. Despite all the efforts made by food organizations to combat the issue, the numbers of people in need stays alarmingly high. The goal of this project is to understand how to effectively connect with the Hispanic community (and potentially other communities in need) through the social media and other information systems. Particularly, this research aims to understand the linguistics and emotional tonality used in the social media by the food organizations.

Background

The field of automatic text generation using AI has advanced incredibly. We surveyed, compared, and analyzed several Artificial Intelligence four Text Generators. Here we focus for of them available online: OpenAI, Rytr.me, CopyAI, and Writesonic. Each of these text generators, to some extent, can generate text based on different human emotions. All the text generators mentioned above are based on GPT-3 (Generative Pre-trained Transformer 3), which is an autoregressive language model that uses deep learning to produce human-like text.

Methodology/Approach

- **Survey**: We surveyed several text generator tools. The selection criteria is based on their abilities to: generate understandable freestyle text, set the tonality of the text, rewrite the text with a different tonality, develop on the previously generated text, create various content (tweets, Facebook posts, ads, essay paragraphs, etc.), and write or translate in a language other than English.
- **Select Tools**: We selected four AI-based text generators: OpenAI API, Rytr.me, CopyAI, and Writesonic.
- **Data Collection**: To collect the data from the text generators, we "fed" them with the keywords and sentences. We used the same input expressions for all four text generators. The keywords, sentences, and phrases were: "Food insecurity in the United States."; "Hispanic population of NYC and Chicago."; "Food pantries"; "Food providing organizations"; "SNAP".
- **Free Style:** Then we generated a freestyle text, not loaded with any emotion or tonality. After proofreading multiple samples of the generated text, we were able to choose the most understandable and user-friendly text.
- **Tonality**: We set the different tonalities to the text-generating tools. The tonalities we chose to experiment with were: "Professional", Compassionate", and "Convincing". We then repeated the process of selecting data now with the tone-based text.
- **Text Analysis:** We used the Bag of Words model to analyze the lexicon used by each text generator. We got rid of stop words, such as "the", "that" "is", etc., and scan the text for repetition.
- **Similarity Analysis**: We calculated the Jaccard **similarity among the lexicon used by each tool**. The Jaccard similarity between two data sets can be written as follows:

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|}$$

where $|A \cap B|$ gives the number of members shared between both sets and $|A \cup B|$ gives the total number of members in both sets (shared and un-shared).

Results/Discussion

Free Style: The highest performance when working with a freestyle text was shown by OpenAI API.

Tone Based: When working with tone-based text, the best outputs were shown by Writesonic and CopyAI.

Word Frequency (Bag of Words and Word clouds)

- Word clouds are images composed of words used in a text. The size of each words represents its frequency or importance.

Figure 1: Word clouds from the lexicon of each text generation tool

- The words "Hispanic", "food", and "insecurity" are the most noticeable.
- OpenAI, CopyAI, and Rytr.me, provided general general statements and statistics from the Web.
- Writesonic was able to build upon the keywords, and provide the reasons behind the food insecurity and ways to battle it.
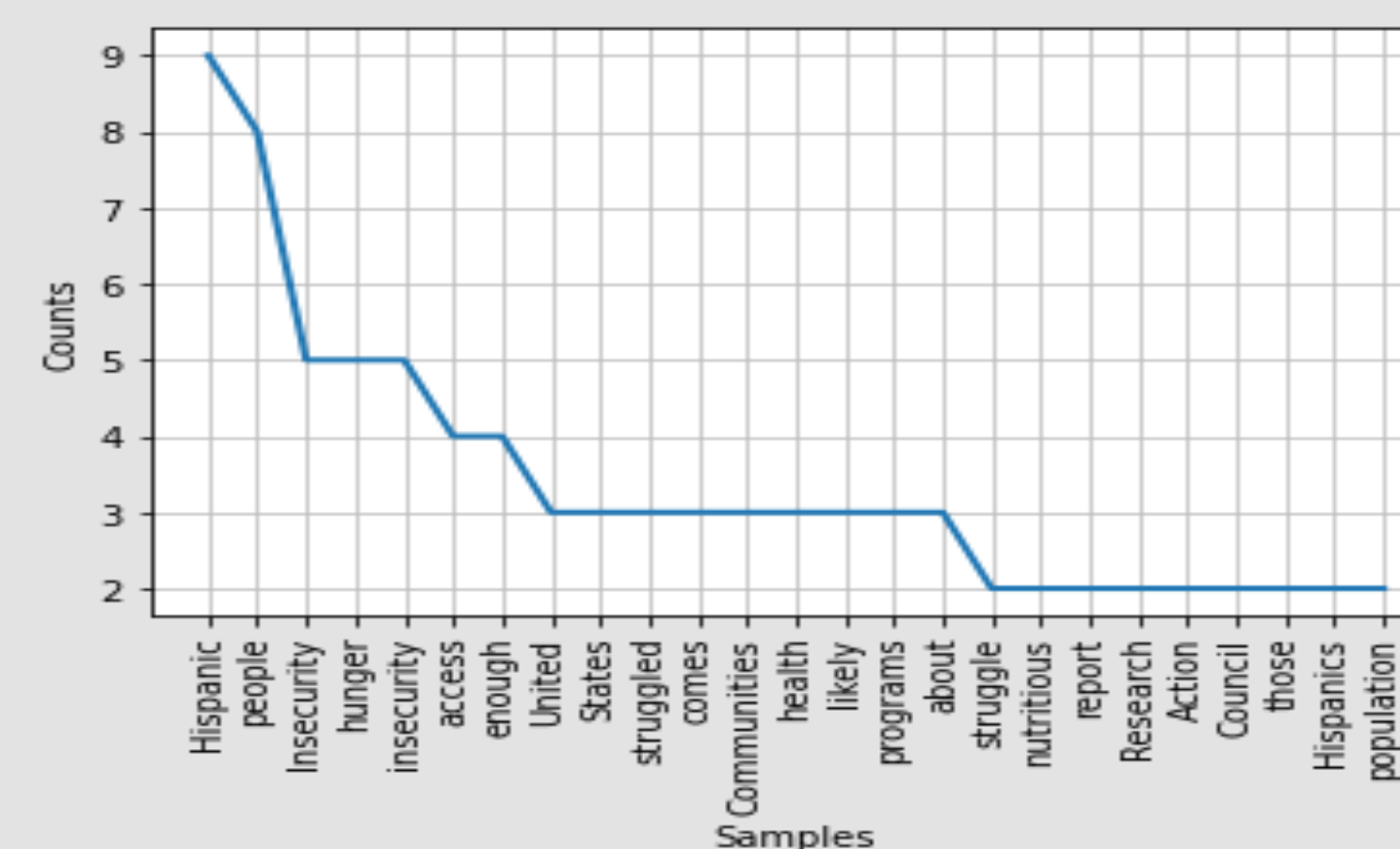

Fig. 2. Bag of Words Model representation in the form of the plot.

The plot on the right is representing the text sample taken from **Writesonic** in a freestyle-professional format.

Analysis of the similarity

Pairs of Text Generators	Text Tonality			
	Freestyle	Professional	Compassionate	Convincing
OpenAI API & Rytr.me	44	N/A	N/A	N/A
OpenAI API & CopyAI	39	N/A	N/A	N/A
OpenAI API & Writesonic	41	N/A	N/A	N/A
Rytr.me & CopyAI	37	16	17	18
Rytr.me & Writesonic	39	17	18	19
CopyAI & Writesonic	43	22	23	24

Figure 3: Jaccard Similarity Between Each Pair of Tools

Results/Discussion (Continue)

Figure 4: Jaccard similarity Charts

a) Freestyle Text

b) Tone-based Text
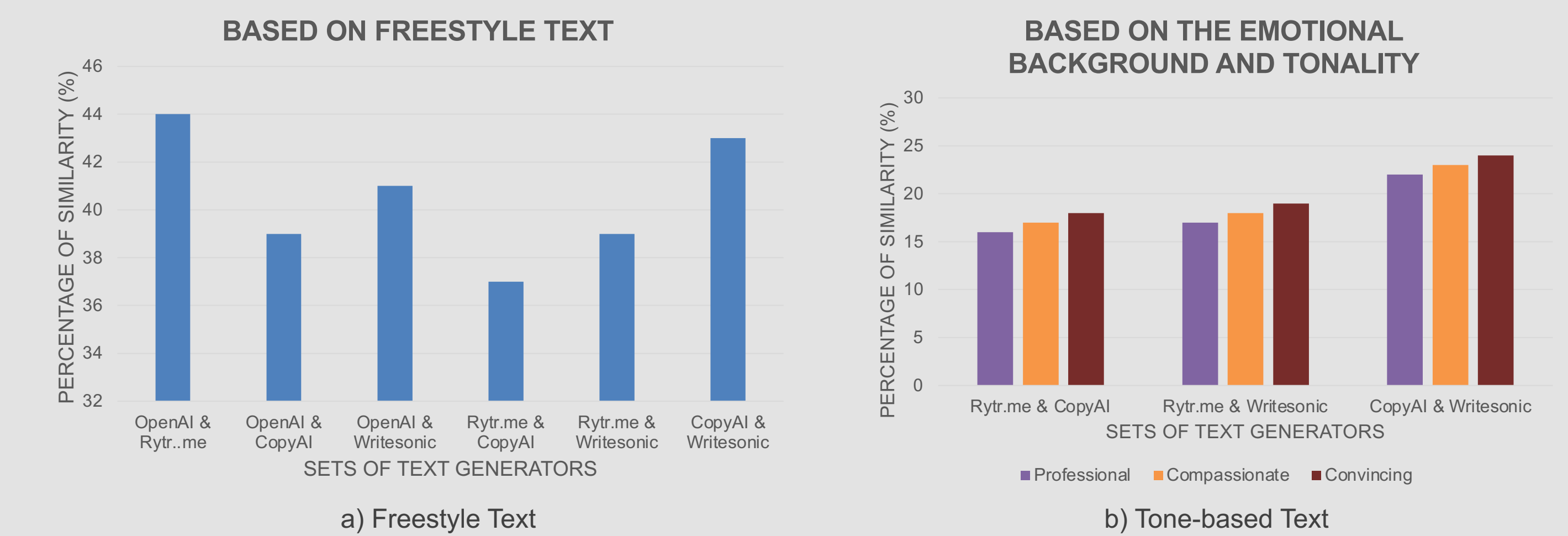
- From Figures 3 and 4 we can see that OpenAI API and Rytr.me have the highest similarity (44%)
- Closely followed by the CopyAI and Writesonic (43%), these two provide similar features such as writing outlines, articles, bullet points, etc.
- OpenAI API and Writesonic (41%) have the third highest similarity.
- In Figure 4, Chart b), we can see the same patterns among tone-based text comparisons as freestyle text.
- We observed that OpenAI API and Rytr.me share a simple design in their the platforms, therefore produce somewhat similar text.

Conclusion

After performing the survey, analysis and comparison of the four text generators available on the market, we concluded that Writesonic outperformed the other three.

All four text generators can generate human like text. However, while OpenAI has the best sentence structure and grammar out of four, we simply cannot set the tonality of the generated text. Rytr.me lacks the grammar skills compared to the other three, and when generating larger amounts of text noticed to be repetitive. CopyAI has both ability to set the tonality of the text and outstanding sentence structure, however this text generator is not capable of writing in the format of essay. The generated text is usually not longer than one paragraph. CopyAI is the tool more suitable for commercials, business presentation outlines, etc. Writesonic combines the grammatical competence of OpenAI API, the functionality of CopyAI, and the ability to set different text tones like Rytr.me. Therefore, this text generator is the best suited option for the main goal of the project – increasing engagement rate with the communities.

Future Work

- Perform the survey among Hispanic population of NYC, Chicago, and LA in order to understand is the text generated by Artificial Intelligence is likable, understandable, and popular.
- Work on understanding the principle of GPT-3 and soon to be announced GPT-4 models.
- Build upon existing GPT models to create our own text generator dedicated to connect with the food insecure communities.

Acknowledgements

References