Cameron Yenche
DATASCI 266
Professor Bhattacharyya
April 14, 2024

Automating Style-Free Text Generation for Conversational AI: Challenges and Pathways Forward

Abstract
This paper explores the development of a process for generating style-free versions of text, an essential component for achieving few-shot text style transfer as demonstrated by a cutting edge text style transfer model developed by AWS AI Labs (Roy et al.,). Within the context of the paper, this the task of style-free text generation was delegated to a pre-trained LLM, but had to be verified using human supervision. Within this paper, we have taken on the task of building a dedicated model that generates a style-free version of the text [1] [2], thus allowing for potential end-to-end few-shot style transfer without supervision. By leveraging a novel methodology that combines summarization, paraphrasing, and vocabulary simplification techniques we aim to automate the creation of style-free text. This process is crucial for enhancing the adaptability and effectiveness of conversational AI, making interactions more relatable and user-friendly. Our findings indicate a promising direction for future work in achieving more accessible and effective communication through AI-driven chatbots.

Introduction
In the evolving landscape of NLP, chatbots have become pivotal in enhancing user experience across digital platforms. Motivated by obvious market demands, AWS AI Labs recently introduced a few-shot model for text style transfer, signaling a significant leap forward from previous models (Madann et al.,). This model aims to make chatbot interactions not only more engaging but also more relatable to users by adapting the conversation style to match that of the user. Such a capability underscores the potential of chatbots to transcend their current functionalities, promising a future where they can understand and mimic human conversational nuances more effectively.

At the core of this advancement lies the concept of style-free text generation. As mentioned, the AWS AI lab paper used an LLM to generate style-free text. However, the scalability and efficiency of this process are hampered by its reliance on human verification. Recognizing this bottleneck, our project focuses on automating the generation of style-free text. By doing so, we aim to facilitate a more streamlined and flexible approach to text style transfer, enhancing the adaptability and effectiveness of chatbots. By enabling chatbots to communicate in a manner that is both more relatable and contextually appropriate, we foresee a significant increase in their usage and effectiveness. This paper outlines our methodology, experiments, and findings in pursuit of this goal, offering insights into the challenges and potential of fully automating style-free text generation.

---

[1] While the paper by Roy et.al describes using an pre-trained LLM to generate style-free versions of the text, but notes that it is a difficult that requires knowledge of the target style (on page 122).

[2] Style-free text conversion is also described by *Madann et al.,* but was not few-shot

Related Work
While there are plenty of papers on the topic of text style transfer, the use of style-free text as an intermediate step is a relatively novel approach[3]. We have two potential hypotheses for why this may be the case:
1) Style-free text can be, in itself, considered a type of styled text.
2) The creation of Style-free text is not a unique task, but shares similarities with summarization and paraphrasing tasks.
While it is not within the scope of this work to adequately explore our first hypothesis, we have taken to testing the second hypothesis in such that it informed our approach towards generating style-free text.

Dataset Description
Our exploration into automated style-free text generation leverages two principal datasets, one chosen as data to test our model on and the other used during the implementation of vocabulary simplification.

Customer Support on Twitter Dataset[4]

This dataset encompasses a large corpus of customer service tweets and replies. It includes 3 million tweets and replies involving some of the biggest brands on Twitter, such as Apple, Amazon, and Uber. The diverse range of customer interactions contained within this dataset offers a realistic glimpse into the types of conversations commercial chatbots are likely to encounter, making it an valuable resource for our project.

Subtlex US Corpus[5]

The Subtlex US is a corpus derived from movie subtitles that contains approximately 51 million words and serves as a comprehensive representation of the English language. This dataset's unique composition allows us to formulate a representative word frequency map, that we can use to aid in our implementation of vocabulary simplification. This approach operates from the assumption that more common language is simpler, more generic language.

Methodology
The primary objective of this project was to automate the generation of style-free text to facilitate few-shot text style transfer. Our approach combined several NLP techniques, including summarization, paraphrasing, simplification, and synonym replacement, leveraging the T5 and BERT models. This section details the approaches tested in our experiments, including dataset preparation, model selection, and evaluation metrics.

Frequency Dictionary Creation

To facilitate vocabulary simplification, we first generated a frequency dictionary from the Subtlex US corpus. This involved normalizing and tokenizing the corpus text, then calculating word frequencies to identify simpler synonyms for complex words.

Initial Model Configuration

We began with a basic model structure, utilizing the t5-small, t5-base, and t5-large models to experiment with text summarization and simplification tasks. This initial phase aimed to determine the most effective T5 model size.

Intermediate Model Configurations

Building upon our initial experiments, we tested intermediate model configurations incorporating summarization (‘summarize’), paraphrasing (‘paraphrase’), and simplification

---

[3] Style-free text as an intermediate step was pioneered in the paper by Madann et al., and built upon by Roy et al.

[4] Link to Dataset can be found here: https://www.kaggle.com/datasets/thoughtvector/customer-support-on-twitter

[5] Link to Dataset can be found here: http://www.lexique.org/?page_id=241

('simplify') tasks in varying sequences to identify the optimal processing flow, with t5-base identified as the most effective T5 model.

Advanced Model Configuration

Our advanced model configuration introduced an additional layer of complexity by incorporating vocabulary simplification based on word frequency. This involved identifying synonyms for complex words that appeared less frequently in the Subtlex US corpus and replacing them with simpler, more common alternatives. This process was iterated twice for each text to ensure thorough simplification.

Evaluation Metrics

The effectiveness of each model configuration was evaluated using several metrics:

- **BERTScore:** Assessed the semantic similarity between the original and transformed texts, ensuring the preservation of meaning.
- **Flesch Reading-Ease Score:** Evaluated the readability of the transformed text, aiming for a balance between simplicity and clarity.
- **Lexical Richness Metrics:** Included the Type-Token Ratio (TTR) and Measure of Textual Lexical Diversity (MTLD) to evaluate the diversity and complexity of the vocabulary used in the transformed texts. These metrics were evaluated as a proxy for style, where highly styled text is taken to have high lexical diversity.

Experimentation and Results

Our experimentation involved applying the developed models to a subset of the Twitter Customer Support dataset, processing the texts in batches to manage computational load. The results were then analyzed to compare the performance of different model configurations and identify the most effective approach for generating style-free text.

Results

Our experimental investigations aimed to automate the generation of style-free text through a series of methodological steps and model configurations. Here are the key findings from our experiments:

Performance Metrics Overview[6]

[6] See next page for figure

Figure 1: The above figure depicts the results from each of the model variations tested using the validation data.

As you can see from *figure 1*, there is a tradeoff between maintaining BERT Score Metrics, decreasing lexical diversity, and increasing readability. Based on our model selection, we were able to achieve significant increases in readability, while decreasing lexical diversity and suffering only slight decreases in BERT score metrics.

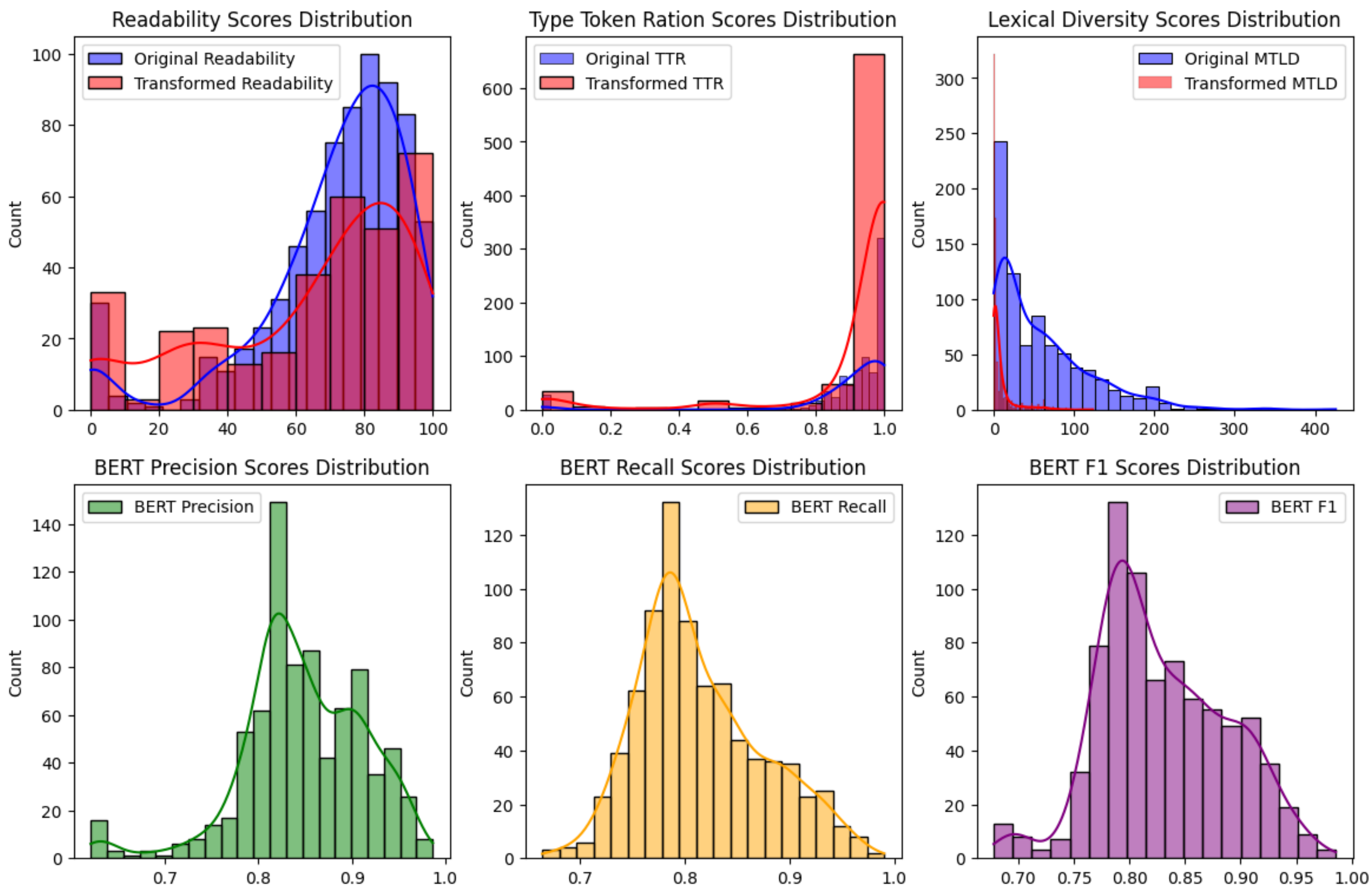

Figure 2: This figure above depicts the distribution of the metrics scores tracked for the final model applied to the test dataset. As you can see, there is a notable increase in readability and decrease in lexical diversity. Note that our BERTScores bins contain no values close to 0.

As you can see from the plot above, our model performed as expected when moving from the validation dataset to the test dataset. We were able to, at large, increase readability, decrease lexical diversity metrics, while maintaining relatively high BERTScore metrics. Important to note here is that there were throw-away values for the readability score metric (i.e., values outside the expected bounds of 0-100), we suspect this could be attributed to more extreme internet vernacular, which may emojis. Even more important to note is that none of the BERTScore metrics have fallen below 0.5, which implies that the model maintains relative cohesion with the original message content in every text in the test dataset (i.e., there were no spectacular failures).

Discussion

This project embarked on the ambitious task of automating the generation of style-free text, aiming to enhance the versatility and effectiveness of conversational AI through few-shot style transfer. The exploration through various model configurations and approaches has not only demonstrated the feasibility of this objective but also illuminated the complexities inherent in text simplification and style-free text generation. Here are the main points of discussion arising from our work:

Balancing Simplicity with Semantic Integrity and Lexical Richness

One of the most striking findings from our experiments was the delicate balance required between simplifying text to enhance readability and maintaining the semantic integrity and lexical richness of the original text. Many of our models enhanced text simplification to such an extent that it harmed readability and cohesion with the original intent of the text.

The Role of Model Configuration in Text Transformation

The comparison between different T5 model sizes and configurations underscored the significant impact of model choice on the success of style-free text generation, where the T5-base model provided more balanced outputs. This finding suggests that while larger models may offer more sophisticated language capabilities, they do not necessarily result in better outcomes for this specific task. Instead, the integration of targeted simplification processes appears more crucial. It is important to note, that we did go down the path of fine-tuning our own model to better accomplish the task of vocabulary simplification by training on the Simple English Wikipedia dump[7], but computational resource constraints prevented the model from being trained within the confines of this project.

The Importance of Vocabulary Simplification

Our approach to vocabulary simplification—replacing less common words with their more frequent synonyms—was a key innovation of this project. After all, a large component of a text's style is its diction. However, the cohesion with the original meaning of the text proved difficult to maintain, despite our efforts to pair word frequency mapping with cosine similarity. This aspect is particularly relevant in contexts where the richness of language plays a crucial role in conveying nuanced meanings or maintaining user engagement. While we made significant stride towards generated simplified vocabulary, this is certainly an arena of the project that would benefit from further work.

Conclusion

This project has taken a crucial step towards integrating style transfer within modern commercial chatbots by automating style-free text generation. Despite our progress, challenges remain that highlight the need for further development with respect to the following:

- **Optimization:** The model's processing time of 4-7 seconds exceeds user-friendly thresholds for chatbots, which highlights the need for efficiency major improvements prior to deployment.
- **Multilingual and Emoji Support:** Our current model's limitation in handling non-English text and emojis—a staple of modern digital communication—presents a crucial area for enhancement. To this end future work may explore integration with multimodal models and more complex LLMs.
- **Alternative Models:** Investigating alternative model structures, such as fine-tuning LLMs on simplified text, offers a promising direction for overcoming existing limitations.

Moving forward, addressing these challenges will be key to developing more responsive and tailored chatbots, for a more inclusive future of conversational AI.

[7] Link to Dataset can be found here: https://www.loc.gov/item/2019205402/

References

Aman Madaan, Amrith Setlur, Tanmay Parekh, Barn- abás Poczós, Graham Neubig, Yiming Yang, Ruslan Salakhutdinov, Alan W Black, and Shrimai Prabhu- moye. 2020. Politeness transfer: A tag and generate approach. In Proceedings of the 58th Annual Meet- ing of the Association for Computational Linguistics, pages 1869–1881.

Roy, S., Shu, R., Pappas, N., Mansimov, E., Zhang, Y., Mansour, S., & Roth, D. (2023, November 1). *Conversation Style Transfer using Few-Shot Learning*. ACL Anthology - ACL Anthology. https://aclanthology.org/2023.ijcnlp-main.9.pdf