The 35th Conference on Computational Linguistics and Speech Processing (ROCLING 2023)
Taipei City, Taiwan, October 20-21, 2023. The Association for Computational Linguistics and Chinese Language Processing

A Comparative Study of Generative Pre-trained Transformer-based Models for Chinese Slogan Generation of Crowdfunding

Yu-Cheng Liang
Department of Data Science,
Soochow University, Taiwan
justinlin579@gmail.com

Meng-Heng Zheng
Department of Data Science,
Soochow University, Taiwan
zxcvbnmk1218@gmail.com

Jheng-Long Wu
Department of Data Science,
Soochow University, Taiwan
jlwu@gm.scu.edu.tw

Abstract

In recent years, language generation models have made significant progress and garnered extensive attention, aiming to generate diverse sentences across domains. However, effectively conveying deep semantics within constrained word limits and expressive formats remains a challenging endeavor. Therefore, we utilize GPT-2, GPT-3.5, and Bloom to generate slogan. Incorporating product descriptions, we have experimented, using metrics like ROUGE, BLEU, and semantic relevance for model evaluation. Overall, compared to product descriptions, GPT-3 demonstrates the best similarity in terms of vocabulary and meaning. In terms of human evaluation results, Bloom better captures the uniqueness of the slogan, while GPT-3 is more closely related to the description, and its sentences are the most fluent.

Keywords: Chinese Slogan Generation, Pre-trained Models, Crowdfunding, Semantic Similarity

1 Introduction

A Slogan, as a part of advertising copy, serves the purpose of attracting more sales or visits and conveying the unique features of a product or service within concise sentence. Some studies have confirmed that slogans on products can enhance consumer recognition of specific goods or businesses, reducing uncertainty in their shopping decisions. With consumers exposed to thousands of advertisements daily, designing a concise and impactful slogan becomes paramount.

Designing a deeply ingrained slogan is not an easy task, and it can also be quite costly. Research by Dimofte and Yalch (2007) indicates that the estimated cost of developing an effective slogan could reach up to one million dollars, yet there is no guarantee of success.

In previous works on slogan generation, most studies primarily employed the sequence-to-sequence (seq2seq) transformer model. However, we observed that when encountering longer input sequences, the seq2seq model fails to capture the entire context adequately, leading to less fluent or inaccurate sentence generation. Therefore, to gain a comprehensive understanding of the product description for the purpose of effective slogan generation, three distinct generative models were employed: GPT-2, GPT-3.5 and Bloom. To ensure consistency with product descriptions, we experimented with incorporating product descriptions into the training process and conducted comparative analyses, including ROUGE, BLEU and semantic relevance.

The key contributions of our research are outlined as follows:

- Assessing the feasibility of Chinese commercial slogan generation.
- Comparing the performance among GPT-2, GPT-3.5 and Bloom models in Chinese commercial slogan generation.
- Offering commercial slogan design strategies.

2 Related Work

In the domain of Natural Language Processing (NLP), Natural Language Generation (NLG) has played a pivotal role in transforming non-linguistic data into human-like text (Reiter et al., 1997). This section offers a succinct overview of this evolution and outlines the rationale for model selection, setting the stage for the subsequent exploration of slogan generation.

Traditional neural networks laid the groundwork for early text generation, emphasizing neuron behavior, connections, and learning (Auli et al., 2013). Convolutional neural networks (CNNs) emerged to capture intricate features through layered structures (Gu et al., 2018). Later, recurrent neural networks (RNNs) excelled in handling sequential data for generating context-rich outputs (Sutskever et al., 2011).

The introduction of generative adversarial networks (GANs) marked a significant breakthrough in text generation by leveraging adversarial training (Crewswell et al, 2018). This approach significantly contributed to diverse and authentic textual outputs. Transformer architecture further revolutionized the field by capturing overarching dependencies within sequences (Keskar et al., 2019).

While extensive research has enhanced text generation across languages, the realm of slogan generation, especially in Chinese, remains understudied. In comparison to languages like English (Tomašic et al., 2014) and Japanese (Iwama et al., 2018), Chinese slogans have received limited attention. This study addresses this gap by exploring Chinese slogan generation using three models: Bloom, GPT-2, and GPT-3. We aim to not only understand their respective capabilities in generating impactful Chinese slogans but also to compare their effectiveness. Our choice of these models stems from the desire to comprehend the capabilities of the newly introduced traditional Chinese pre-trained model, Bloom-zh, and to benchmark against well-known pre-trained models like GPT-2 and GPT-3.

3 Methodology

First, data collection content is defined and executed. Following data acquisition, model training is carried out, and different input data are categorized based on experimental purposes. Continuous fine-tuning of the model occurs during the training process. Finally, the model training outcomes are evaluated using evaluation metrics to assess the effectiveness of the model. The research process is depicted in Figure 1:

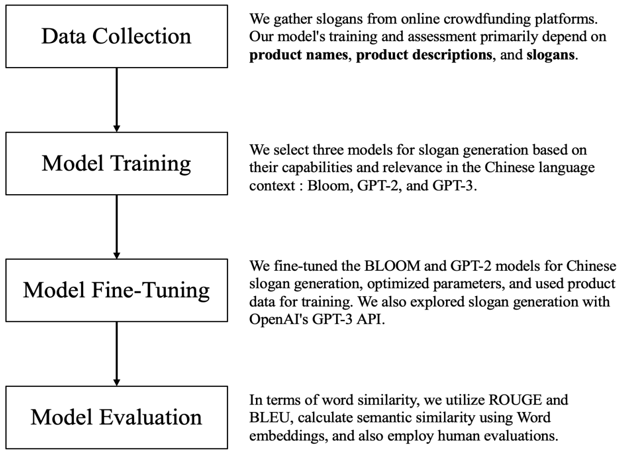

Figure 1: The Research Process.

3.1 Data Collection

We crawl commercial slogans from online crowdfunding platforms such as flyingV[1], WaBay[2] and zeczec[3]. The reasons for using online crowdfunding platforms as a data collection method are as follows:

- **Diverse Range of Slogans:** Our dataset spans across technology, education, gaming, music, and more.
- **Fresh and Up-to-date Content:** New crowdfunding projects on the platform continue to emerge, suggesting the dataset is likely up to date, reflecting current trends and market preferences.

We organized the scraped data according to the column format in Table 1. Column 1 represents the field name, and Column 2 provides an explanation for that field. Due to variations across platforms, some fields may have missing data. Additionally, we define product descriptions as **concise descriptions that combine product functions and features etc**. Among the three platforms, only zeczec provides defined product descriptions.

Apart from data directly sourced from crowdfunding platforms, certain slogans may exist within images rather than textual data. Hence, we have developed a GUI interface to assist us in manually extracting slogans and product names.

[1] https://www.flyingv.cc/

[2] https://wabay.tw/

[3] https://www.zeczec.com/

Name	Description
name	Product Name
proposer	Proposer Name
projLink	Project Link
imgLink	Product Image Link
projType	Project Type
raised	Amount Raised
aim	Planned Funding Goal
numPeople	Total Number of Backers
completed	Goal Achievement
crawltime	Crawling Time
dataFrom	Data Source
prodType	Product Type
projPerc	Project Success Rate
starttime	Campaign Start Time
endtime	Campaign End Time
descr	Product Description

Table 1: Column Form

Our model's training and evaluation primarily rely on **product names**, **product descriptions**, and **slogans**.

3.2 Model Selection

To fulfill our research objectives, we meticulously opted for three distinctive models to facilitate the task of slogan generation: Bloom, GPT-2, and GPT-3. The rationale behind these choices was underpinned by their inherent capabilities and relevance within the Chinese language milieu. In particular, the recently introduced traditional Chinese pre-trained model, Bloom-zh, was harnessed to explore its prowess in crafting impactful Chinese slogans. Moreover, GPT-2 and GPT-3 were embraced as benchmark models due to their well-established proficiency in text generation endeavors.

3.3 BLOOM Fine-tuning

Configuration and Setup: We initiated the fine-tuning process by loading the Bloom-zh model from the "ckip-joint/bloom-1b1-zh" pre-trained checkpoint. The model was instantiated using the "AutoModelForCausalLM" class, and its tokenizer was loaded using the "AutoTokenizer" class from the Hugging Face Transformers library. Model parameters were optimized, and specific parameters were cast to 'torch.float32'.

PEFT Model Enhancement: To enhance the model's performance, we introduced Lora, a Parameter Efficient Fine-Tuning (PEFT) method. The "LoraConfig" was tailored with parameters like "r" and "lora_alpha" targeting specific model modules such as "query_key_value." The PEFT-enhanced model was obtained using the "get_peft_model" function from the "peft" module.

Data Preparation: We utilized a dataset comprising product names, descriptions, and slogans. A prompt generation function was designed to combine these elements, which were then tokenized using the pre-loaded tokenizer.

Model Training: The transformed dataset was used for fine-tuning the model using the "transformers.Trainer" class. Training arguments were configured with batch sizes, learning rates, and other hyperparameters. The model underwent training using the "trainer.train()" method.

Inference and Slogan Generation: For slogan generation, an inference function was devised that generated prompts based on product names. The model was utilized to generate corresponding slogans, and the output was decoded using the tokenizer.

3.4 GPT-2 Fine-tuning

Configuration and Setup: We configured the GPT-2 model for slogan generation using the Hugging Face Transformers library. The "uer/gpt2-chinese-cluecorpussmall" pre-trained model, tailored for Chinese language tasks, was employed.

Tokenization and Special Tokens: Tokenization was facilitated using the BertTokenizer, which introduced special tokens like '<name>', '<description>', and '<slogan>'. These tokens segregated different input segments, representing the product name, description, and slogan.

Embedding Adjustment: Model embeddings were resized to accommodate the new special tokens, ensuring efficient processing of the modified input data.

Dataset Preparation: Our dataset preparation involved constructing tokenized examples and managing annotations for product names, descriptions, and slogans.

Model Training: Fine-tuning spanned three epochs using the AdamW optimizer. The model was optimized to generate slogans in line with provided product name and description inputs.
Evaluation and Slogan Generation: Post-training, the evaluation stage involved a sampling function designed for slogan generation. This function utilized techniques such as top-k and nucleus (top-p) filtering to govern the slogan generation process, resulting in slogans encapsulating provided inputs.

3.5 GPT-3 API Utilization

Configuration and Setup: We leveraged the "gpt-3.5-turbo" pre-trained model from OpenAI's GPT-3 API to explore Chinese slogan generation.

Dataset Preparation: Our dataset preparation involved concatenating product names and descriptions to construct prompts for the GPT-3 model, enabling slogan generation.

Slogan Generation using GPT-3: For slogan generation, we used the GPT-3 model by providing concatenated product names and descriptions as prompts. We used the default model parameters for generating slogans.

3.6 Evaluation Metrics

ROUGE-N: ROUGE is a set of metrics commonly used for evaluating the quality of machine-generated summaries. ROUGE 2 and ROUGE L have been demonstrated to be suitable for evaluating single document summaries, while ROUGE-1 and ROUGE-L are applicable for very short summaries or headline-like summaries (Lin, 2004). We used the average score of ROUGE-N which focuses on the matching of n-grams, where n represents the length of words. As the data is in Chinese, prior to computation, it's necessary to perform segmentation on the data. We utilize the Jieba package (version 0.42.1) and computed the ROUGE score using the publicly available Python script of rouge (version 1.0.1).

BLEU: BLEU (Bilingual Evaluation Understudy) is a metric used to assess the quality of machine-generated translations. It evaluates the similarity of n-grams between the machine-generated translation and one or more reference translations. Following a similar concept, we employ BLEU to evaluate the lexical similarity between the generated slogans and the original slogans and product descriptions.

Semantic Relevance: Traditional methods use linguistic features like word overlap, n-grams, and syntax for similarity scoring. However, these methods may not capture sentence semantics well. Word embeddings represent words in continuous vectors, enabling sentence comparison using these word vectors. We employed three different transformers: paraphrase-xlm-r-multilingual-v1 [4] and distiluse-base-multilingual-cased-v2 [5] are designed to support multilingual usage, while text2vec-base-chinese[6] is specifically suitable for Chinese characters and words.

3.7 Human Evaluation Methodology

In our study, a single evaluator directly rated the slogans generated by our Slogan generation model using three key criteria: distinctiveness (Distinc.), adequacy (Adeq.), and fluency (Flu.), with a rating scale ranging from 1 (lowest) to 5 (highest).

Distinctiveness: Measures the uniqueness and specialization of each generated slogan.

Adequacy: Evaluates how well each slogan captures the essence of the context or target.

Fluency: Assesses the naturalness and readability of each slogan.

4 Experimental Settings

4.1 Dataset

We have obtained a total of 11,284 records, with 3,491 from FlyingV, 749 from WaBay, and 7,044 from zeczec (including records with product descriptions).

Among them, 7,674 projects have reached their funding goal, 3,346 projects did not meet their funding goal, and 264 projects are categorized as

[4] https://huggingface.co/sentence-transformers/paraphrase-xlm-r-multilingual-v1

[5] https://huggingface.co/sentence-transformers/distiluse-base-multilingual-cased-v2

[6] https://huggingface.co/shibing624/text2vec-base-chinese

long-term sales projects. Based on a grouping of 150% intervals up to 1000%, the majority of data is concentrated within the 100%-250% range, totaling 3,911 records, which approximately accounts for 35.5% of the total dataset. Additionally, the average length of product descriptions is around 65 words.

4.2 Pre-trained Models

This study utilizes three pre-trained Chinese models: GPT-2, GPT-3.5 and Bloom. Among them, both Bloom and GPT-2 undergo fine-tuning before generating, whereas GPT-3.5 generates directly through prompts.

GPT-2: We employed the gpt2-chinese-cluecorpussmall model from uer [7]. Adjustments to default parameters were made as follows:

- **"max_length=20":** This configuration led to slightly shorter average lengths while yielding improved performance.
- **"repetition_penalty=1.2":** Discouraging excessive repetition in generated slogans.
- **"temperature=1":** Speculating that overly high temperatures might deviate from the intended outputs due to insufficient model training.

GPT-3: For GPT-3, we utilized the OpenAI API for direct slogan generation without parameter adjustment.

- prompt='Generate suitable ad slogans based on the given product names and descriptions. Product names and descriptions: [product_name], [product_description] Slogan:'.[8]

Bloom: We conducted experiments on the Bloom-1b1-zh model provided by ckip-joint[9]. Parameter adjustments beyond defaults were as follows:

- **"max_length=35":** Aligning with the average slogan length.
- **"repetition_penalty=1.3":** Discouraging excessive repetition in generated slogans.
- **"temperature=1.3":** Encouraging more creative and diverse output in slogan generation.
- prompt=' Product names and descriptions: [product_name], [product_description] Slogan:'.[10]

In addition to incorporating the original slogans of the products into the training, we also attempted to include the product descriptions, aiming to observe their effectiveness and impact.

5 Results and Case Study

5.1 Word Similarity

Table 2 and Table 3 illustrate the results of automated evaluation, both presenting the average

	GPT2	GPT3	BLOOM
Without product descriptions			
ROUGE 1	0.0589	0.1302	0.1422
ROUGE L	0.0543	0.1145	0.1243
BLEU	0.0067	0.1817	0.0182
With product descriptions			
ROUGE 1	0.0415	0.1700	0.0885
ROUGE L	0.0385	0.1448	0.0796
BLEU	0.0058	0.2369	0.0125

Table 2: Automatic evaluation results. Compare with the original slogan.

	GPT2	GPT3	BLOOM
Without product descriptions			
ROUGE 2	0.0013	0.0083	0.0136
ROUGE L	0.0358	0.0780	0.0907
BLEU	0.0000	0.0002	0.0002
With product descriptions			
ROUGE 2	0.0006	0.0223	0.0080
ROUGE L	0.0250	0.1063	0.0689
BLEU	0.0000	0.0020	0.0002

Table 3: Automatic evaluation results. Compare with the product description.

[7] https://huggingface.co/uer/gpt2-chinese-cluecorpussmall

[8] prompt = '根據各個商品名稱與描述給出適合的廣告詞。商品名稱與描述:{name, description}廣告詞:'

[9] https://huggingface.co/ckip-joint/bloom-1b1-zh

[10] prompt = '商品名稱和描述:{name}:{ description}廣告詞:'

scores across the entire dataset. Note that slogans involve creativity, resulting in comparatively lower ROUGE score compared to other tasks.

The two tables below are both divided into two sections based on whether product descriptions were included during the training process. The scores correspond to the models listed above and the evaluation metrics on the left. Taking Table 2 as an example, when GPT-2 was trained without including product descriptions, the generated slogans were evaluated using the ROUGE-1 metric, resulting in a score of 0.0589.

ROUGE: The results indicates that regardless of whether there was an addition of product descriptions during training, the performance gap between GPT-2 and other models was significant.

As shown in the table, we can notice that when the model is not trained with product descriptions, the bloom model shows higher correlation with original slogans, while incorporating product descriptions into training. Results in GPT-3 exhibiting higher relevance. One reason for this may be that when incorporating product descriptions into training, the BLOOM model may experience a decrease in relevance when generating new slogans due to potential overfitting to the original data. On the other hand, GPT-3 has the advantage of being trained on a wider variety of text types, resulting in better generalization. As a result, even in the presence of product descriptions, GPT-3 can generate slogans with higher relevance.

BLEU: Unlike the ROUGE scores, GPT-3's BLEU scores consistently surpass those of the other two models, whether product descriptions are included in the training process. The differences in score results may be attributed to the fact that ROUGE emphasizes recall, while BLEU places more emphasis on precision.

5.2 Semantic Relevance

Table 4 and Table 5 illustrate the results of evaluating word semantic relevance and calculate the average scores across the entire dataset.

The contents of these two tables are separated by three different transformers, with the scores corresponding to the models listed above and the relationships to be compared on the left side: "**Des**" represents **product descriptions**, "**OG**" represents **original slogans**, and "**New**" represents **slogans newly generated by different models**. Taking Table 4 as an example, we observe a score of 0.2837, which corresponds to the semantic relationship score between the product description (Des) and the slogan newly generated (New) by the GPT-2 model without incorporating product descriptions training. This score is obtained within the transformer of "paraphrase-xlm-r-multilingual-v1".

	GPT2	GPT3	BLOOM
paraphrase-xlm-r-multilingual-v1			
Des & OG	0.4675	0.4675	0.4675
Des & New	0.2837	0.4752	0.4034
New & OG	0.2964	0.4606	0.4165
distiluse-base-multilingual-cased-v2			
Des & OG	0.2799	0.2799	0.2799
Des & New	0.0832	0.2908	0.2374
New & OG	0.1617	0.2926	0.2618
text2vec-base-chinese			
Des & OG	0.6486	0.6486	0.6486
Des & New	0.5537	0.6675	0.6162
New & OG	0.4926	0.5958	0.5715

Table 4: Sentence similarity evaluation results. Training **without** product descriptions.

	GPT2	GPT3	BLOOM
paraphrase-xlm-r-multilingual-v1			
Des & OG	0.4675	0.4675	0.4675
Des & New	0.3214	0.6752	0.4354
New & OG	0.3533	0.4961	0.4039
distiluse-base-multilingual-cased-v2			
Des & OG	0.2799	0.2799	0.2799
Des & New	0.1018	0.5553	0.2756
New & OG	0.1727	0.3201	0.2470
text2vec-base-chinese			
Des & OG	0.6486	0.6486	0.6486
Des & New	0.5770	0.8086	0.6400
New & OG	0.4960	0.6385	0.5657

Table 5: Sentence similarity evaluation results. Training **with** product descriptions.

In the evaluation of word semantics, GPT-3 shows a high degree of relevance among product descriptions, original labels, and newly generated labels, regardless of whether product descriptions are included in the training. After incorporating product descriptions into training, there is a noticeable improvement in word semantic relevance, and the newly generated labels are more closely aligned with the descriptions of the products compared to the original labels.

On the contrary, BLOOM and GPT-2 do not show a significant improvement in word semantic relevance after incorporating product descriptions into training. The original labels, in comparison to the newly generated labels, even remain closer to the descriptions of the products.

5.3 Human Evaluation Result

	Distinc.	Adeq.	Flu.
GPT-2	2.44	1.96	2.12
GPT-3	3.85	**4.22**	**4.82**
BLOOM	**4.06**	3.94	4.73

Table 6: Human Evaluation Results. **Bold** indicates the best average score.

Table 6 represents the results of manual calculations. We generated 150 slogans by randomly selecting descriptions in the test data. Each worker assessed the generated slogans based on distinctiveness (Distinc.), adequacy (Adeq.), and fluency (Flu.), assigning scores from 1 to 5. The final scores were then averaged.

5.4 Case Study

Table 7 presents the description, original slogan, and generated slogan of example product -Vertical smoky bamboo pen (Vertical 燻竹直立筆). Both the product description and the original slogan emphasize the concepts of “on the tabletop” and “microcosm.” We can also derive the product attribute through the keyword “writing.”

In the slogans generated by GPT-2, we can clearly observe inconsistencies in sentence structure and vocabulary usage in Chinese. Additionally, the meanings of the generated slogans differ from the product description and instead emphasize reading. On the other hand, GPT-3’s generated slogans follow the concept of “writing,” stemming from the product description. After incorporating training with descriptions, these slogans are closer in expression to the original slogans and product descriptions. The focal points highlighted by the BLOOM model are slightly different from the original slogans and product descriptions, with the generated slogans focusing on aspects of life.

Description	
Enjoy the beauty of writing, taste a delightful microcosm on the tabletop. (享受書寫的美好・品味桌上的微風景)	
Original Slogan	
A microcosm that everyone can have on their tabletop. (每個人桌上，都能擁有的微風景)	
Generated Slogan	
GPT-2	Bringing you a reading and collecting experience that gazes up at the era from the workshop. (讓你從工關帶來閱讀 收藏體驗仰視時代)
GPT-3	Highlighting your creativity, the vertical smoky bamboo pen, unique writing experience. (點出你的創意，燻竹直立筆，獨特的書寫體驗)
BLOOM	Make your life more wonderful, inspire you to write more heartfelt moments! Record in the simplest way. (讓你的生活更美好，讓你寫出更多感動！用最簡單的方式記錄)
GPT-2*	The whole person can also become an artistic reading understanding the beauty of disassembling. (整個人也可以後會是藝術式閱讀 懂拆裝美觀)
GPT-3*	Savor the beauty of writing, bringing a microcosm of tranquility to your tabletop. (品味書寫的美好，帶給你桌上的微風景)
BLOOM*	Make life better, write your story in the simplest way possible. (讓生活更美好，用最簡單的方式寫出你的故事)

Table 7: Example outputs - Vertical smoky bamboo pen (Vertical 燻竹直立筆). * indicates model trained with product description.

6 Discussion

GPT-3.5: GPT-3.5 demonstrated the best overall performance in generating slogans. Its slogans were consistently high in quality and contextually relevant, even without fine-tuning. This stability in performance could be attributed to its large model size and diverse training data, which enable it to understand and generate slogans effectively across various contexts.

Bloom: While Bloom's overall performance didn't match that of GPT-3.5, it excelled in generating distinct and creative slogans. This might be due to the specific training data used for Bloom, which included a substantial amount of creative and distinctive slogans. The model's ability to generate creative content could stem from this specialized training.

GPT-2: GPT-2, while a capable model, generated slogans of lower quality compared to both GPT-3.5 and Bloom. Its limitations in generating slogans may be attributed to its smaller model size and less diverse training data compared to GPT-3.5.

In summary, GPT-3.5's superior performance in generating slogans is due to its model size, diverse training data, and inherent capabilities. Bloom's strength lies in generating distinctive and creative slogans, possibly because of its specialized training data. GPT-2, while competent, falls behind due to its smaller model size and less diverse training data in comparison to GPT-3.5.

7 Limitation

Firstly, due to limitations, GPT-3.5 is used here without fine-tuning. After fine-tuning, the results may be more creative. Furthermore, because it is difficult to define what makes a good or bad slogan and considering the uniqueness of slogans, there are only subjective criteria.

8 Conclusion

The purpose of this study is to generate slogans based on the descriptions of target items. After obtaining the dataset from crowdfunding platforms, we applied the original slogans and product descriptions to three slogan generation models: GPT-2, GPT-3, and Bloom. We aimed to evaluate the newly generated slogans from these models.

In terms of word relevance, through ROUGE and BLEU scores, we observed that while BLOOM’s wording closely matches the original slogans, its precision is lower than that of GPT-3. Despite GPT-3 not using words that are very similar to the original slogans or product descriptions, its word choices are more accurate. With the inclusion of product descriptions in training, significant improvements can be seen, especially in GPT-3’s performance.

In terms of word semantics, GPT-3 clearly demonstrates a better grasp of relevant meanings, which is evident in the example outputs as well. In comparison, Bloom slightly deviates from the original product's meaning or description, but not to the extent of GPT-2’s incoherent sentences and incorrect word choices.

9 Future Work

Currently, most research on slogans is based on text data such as product descriptions and attributes to generate slogans. However, product images may potentially assist in generating more accurate and creative slogans. Therefore, we hope to incorporate non-textual data like product images into the model in the future. Additionally, we also aim to increase the diversity of the dataset by exploring sources such as shopping platforms.

Acknowledgments

The research work described in this paper was supported by the grant from the National Science and Technology Council (NSTC), Taiwan, ROC, under Grant No. MOST 111–2221-E-031–004-MY3.

References

Reiter, E., & Dale, R. 1997. *Building applied natural language generation systems*. Natural Language Engineering, 3(1), pages 57-87.

Auli, M., Galley, M., Quirk, C., & Zweig, G. 2013, October. *Joint language and translation modeling with recurrent neural networks*. In Proc. of EMNLP.

Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., ... & Chen, T. 2018. *Recent advances in convolutional neural networks. Pattern recognition*, 77, pages 354-377.

Sutskever, I., Martens, J., & Hinton, G. E. 2011. *Generating text with recurrent neural networks*. In

Proceedings of the 28th international conference on machine learning (ICML-11), pages 1017-1024.

Creswell, A., White, T., Dumoulin, V., Arulkumaran, K., Sengupta, B., & Bharath, A. A. 2018. *Generative adversarial networks: An overview*. IEEE signal processing magazine, 35(1), pages53-65.

Keskar, N. S., McCann, B., Varshney, L. R., Xiong, C., & Socher, R. 2019. *Ctrl: A conditional transformer language model for controllable generation*. arXiv preprint arXiv:1909.05858.

Tomašic, P., Znidaršic, M., & Papa, G. 2014. *Implementation of a slogan generator*. In *Proceedings of 5th International Conference on Computational Creativity, Ljubljana, Slovenia*,Vol. 301, pages 340-343).

Iwama, K., & Kano, Y. 2018, November. *Japanese advertising slogan generator using case frame and word vector*. In *Proceedings of the 11th International Conference on Natural Language Generation*, pages 197-198.

Dimofte, C.V. and Yalch, R.F. 2007. *Consumer Response to Polysemous Brand Slogans*, Journal of Consumer Research, Vol. 33, No. 4 (March 2007), pages 515-522.

Chin-Yew Lin. 2004. *ROUGE: A Package for Automatic Evaluation of Summaries*. In Text Summarization Branches Out, Barcelona, Spain. Association for Computational Linguistics., pages 74-81.

Shotaro Misawa, Yasuhide Miura, Tomoki Taniguchi, and Tomoko Ohkuma. 2020. *Distinctive Slogan Generation with Reconstruction*. In *Proceedings of Workshop on Natural Language Processing in E-Commerce,* pages 87–97.