Pui Kiu College
info@puikiu.edu.hk
Shenzhen College of International Education
info@scie.com.cn
Nanjing University
WEBMASTER@NJU.EDU.CN

Personalized Generative Model via Active Learning
First Edition

Chen Yueming*
Pui Kiu College
Hong Kong, China
email: yuemingdd@gmail.com

Wang Junyi†
Shenzhen College of International Education
Shenzhen, China
email: s21213.wang@stu.scie.com.cn

Dong Yi‡
NANJING UNIVERSITY
Nanjing, China
email: dongyi@nju.edu.cn

11, October, 2023

Abstract

In this paper, we consider such cases when the user did not achieve the preferred result from a deep learning model. A stable diffusion model generates images from a given prompt. For example, a prompt ”Circle” generates an image of a circle, as our expectation. However, if a longer prompt is used, in this case, ”Circle and Triangle”, the model will generate either a circle behind a triangle, or a circumscribed circle, or none of the above. This is not what we want if we expected a picture of circle on top of a triangle. This shows that the reason is the accuracy of the model becomes lower when a longer prompt is used, that our expectation increases when the prompt length increases. In order to tackle this problem, we propose a prediction algorithm to approach the user's preference and generate a suitable phrase to describe the result so that the stable diffusion model would generate similar results. We see it as ”Profitable” since it may be useful for stable diffusion models. In the following article, we will use stable diffusion models as examples. We use a four-choices algorithm to gradually approach the prediction of the user's auction. We understand the picture as a group of n-dimensional coordinates translated by computer. The four choices of pictures help to improve the accuracy of prediction. We code the model to generate images on two axes. When one of the pictures is chosen, the other one on the same axis will be eliminated. Four images on two newly generated axes are generated again and the user has to choose their most preferred image. All the image correction is based on the original image and axis-based picture generated by stable diffusion. The model will stop and have a final predicted prompt for Stable diffusion models to generate user-preferred images. We are planning to create an API for this model. So users are not only able to apply it to Stable diffusion models but also to other models such as the famous GPT model. Currently, we have finished the simple test, verification, and visualization of the model. Image tests are conducted to test the accuracy of the model.

*Pui Kiu College, department of DSE Physics and DSE M1
†Shenzhen College of International Education, Department of A-level English literature, A-level Physics, A-level Math, A-level Chemistry and A-level Biology
‡NANJING UNIVERSITY, Department of Computational Science

Contents

Part I
Introduction

The last decade has witnessed rapid advancements made in machine learning (ML), especially in the deep learning field [3]. One of the popular approaches adopted by many generative models is prompt. Although this method is convenient and straightforward to users, sometimes the prompts provided could be unclear for the AI to understand the users' true preferences [4]. Given that the prompts-driven generative model collects users' preferences from the users only by prompts, we decide to improve this model by enabling the users to make choices (choosing one of the four generated images) until they feel satisfied with the image generated.

Related works In [4], Wang explained that users may opt for "natural expression" as opposed to descriptive language, consequently, the AI may interpret the emotion behind the prompt incorrectly and generate images that completely differ from the user's expectation.

Challenges. The way to get the prediction is hard.

Our Results and Techniques. As for the result, we managed to turn images such as "0" into "1" by the model. This shows that the model is able to transform the image into the target image. For real-life conditions, of course, the target image is only in the user's own choices. The figures show a great example of the result.

Related Works. [1] [2] [4] [5]

The transformation process from 'Initial' to 'Result' by targeting 'Target'

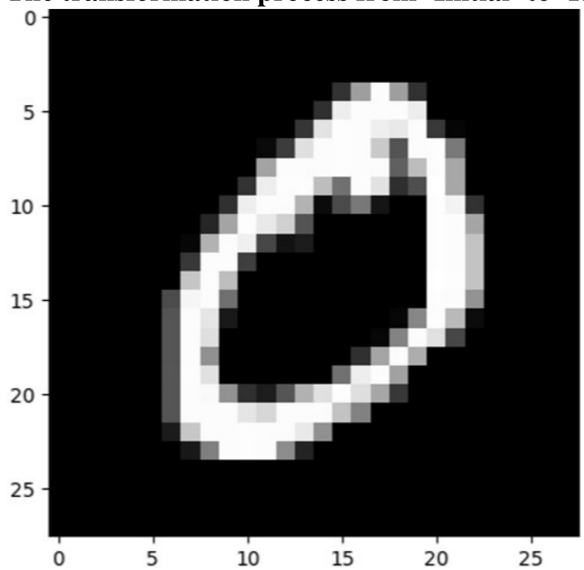


Figure 1: Initial ("0")

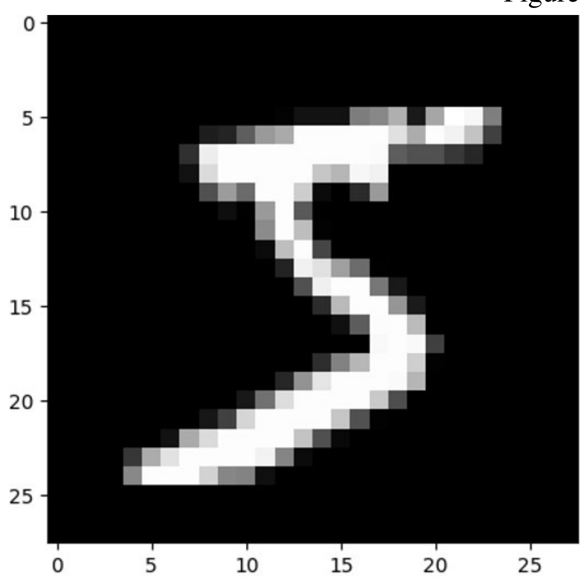


Figure 2: Target ("2")

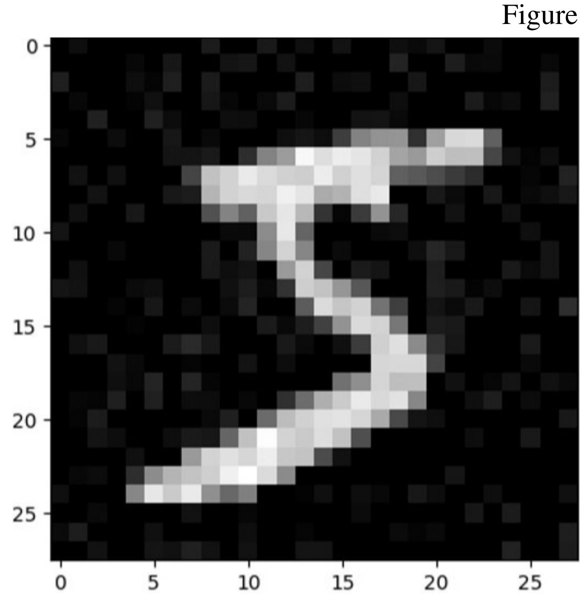


Figure 3: Result ("2")

Part II
Preliminaries

In this section, we formally introduce the background information of how the auction of the user works, ways to improve known results, and how stable diffusion works, as preparation for our machine learning algorithm.

1 Auction of the User

Auction of the user can be categorized into four parts, which are:

- Auction
- Choice
- Expectation
- Want

respectively. Here we will do a brief introduction to the above topics.

1.1 Auction

1.1.1 Basics of Auction

Auction is the process of selling and buying under the usage of bids and private value. There are three types of Auction: 1. Single-item auction

Theorem 1.1 *Single-Item Auction setups with one seller. The seller sells one item to at least one buyer. During the auction, the seller will either maximize the welfare of the auction or maximize the seller's own revenue. Each buyer has his own private value v_i for the object and follows some prior distribution $p_i(v_i)$ where $p_i(v_i)$ is known by the seller while v_i is unknown.*

Example 1 *There is one seller with one item to sell. The seller will try to maximize his own revenue*

2. Second-price auction

Theorem 1.2 *Second-price auction setups with one seller. The buyer observers v_i, while submitting their bids(b_i) simultaneously, where $b_i \neq v_i$. The person with the highest bid wins. The seller charges the winner price which equals the second-highest bid.*

Example 2 *There is one seller with one item to sell. The buyer with the highest bid wins. The seller charges the winner with a price equal to the second-highest bid.*

3. Multi-item auction

Theorem 1.3 *Multi-item Auction setups with one seller. The seller sells different items. There are n agents and a finite set of outcomes. Each agent(i) has a private valuation $v_i(\omega)$ for each outcome where $\omega \in \Omega$.*

Example 3 *There is one seller with multiple items to sell. Each agent has a private valuation for each outcome.*

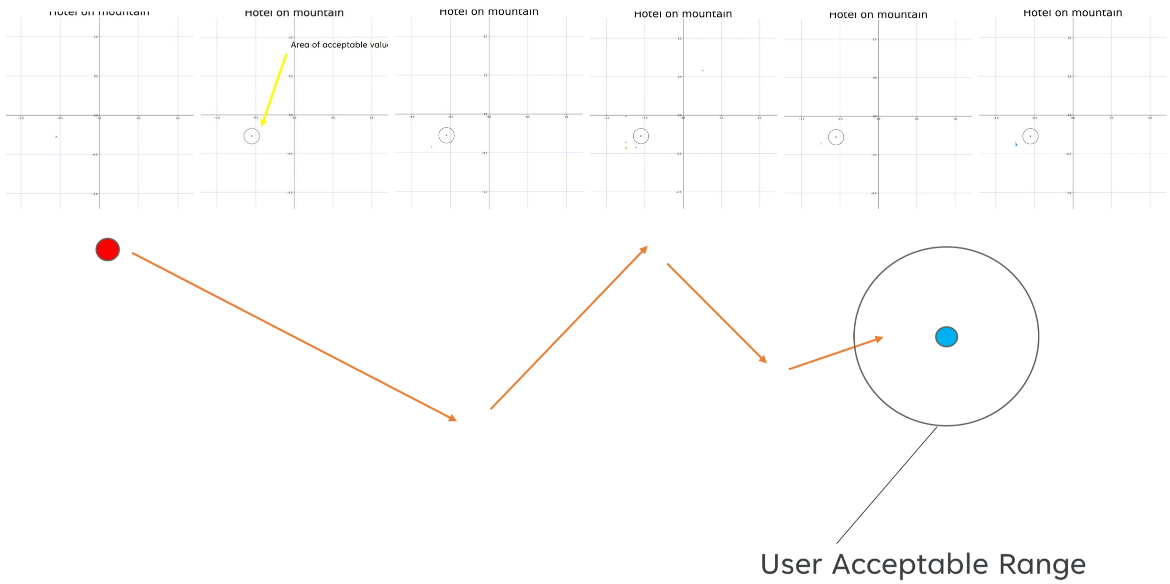


Figure 4: The auction in the picture 'Hotel on the mountain'

1.1.2 History of Auction

During the ancient ages, people bartered for price. This caused the problem of unbalanced cost. In this case, currencies are invented in order to balance the cost. However, for items with changing prices, or undecided prices, auctions were held in order to sell the item at a reasonable price.

1.1.3 Applications of Auction in this algorithm

We can use the idea of an auction in the algorithm
We set the image generated by the stable diffusion model as the base price, such that the target of the algorithm is to raise the base price into the user's price of the auction. As for the visualization, we created the following graphs *Figure 4*.

1.1.4 Formulas Related to Auction

Sender's expectation The following four equations about the sender's expectation: (Shaddin Dughmi, Haifeng Xu, Algorithmic Bayesian Persuasion, 2016)

$$\text{maximize} \sum_{\theta \in \Theta} \sum_{i=1}^{n} \lambda_\theta \varphi(\theta, \sigma_i) s_i(\theta) \tag{1.1.4.0.1}$$

$$\text{subject to} \sum_{i=1}^{n} \varphi(\theta, \sigma_i) = 1, \text{for } \theta \in \Theta \tag{1.1.4.0.2}$$

$$\sum_{\theta \in \Theta} \lambda_\theta \varphi(\theta, \sigma_i) r_i(\theta) \geq \sum_{\theta \in \Theta} \lambda_\theta \varphi(\theta, \sigma_i) r_j(\theta), \text{for } i, j \in [n] \tag{1.1.4.0.3}$$

$$\varphi(\theta, \sigma_i) \geq 0, \text{for } \theta \in \Theta, i \in [n] i \tag{1.1.4.0.4}$$

1.2 Choice

1.2.1 Basics of Choice

Choice is the iteration process of the user selecting the most preferred answer in a bunch of random products.

1.2.2 Applications of Choice in this algorithm

In psychology, there is a term called 'one-third effect'

Definition 1.1 *When there are three choices lying on the same straight line, the user will automatically choose the middle choice.*
This is because users will generate resistance when seeing the choices, but will regret after rejecting the first choices.

This means that for the four generated pictures in our model, the user will have a higher tendency to select the 2^{nd} and 3^{rd} picture.

1.2.3 Cost

Opportunity cost

Definition 1.2 *In the microscopic economy, the opportunity cost is the value of the best forgone where a choice needs to be made between several mutually exclusive alternatives.*
It is the "cost" incurred by not enjoying the benefit that would have been had by taking the second-best available choice.
In general, opportunity cost is the highest-valued option forgone.

Example 4 *Mary used* $\$10,000$ *to buy a flat. What are the opportunity costs of living in it instead of renting it out?*

Example 4 ***(continued)*** *Idea of thinking: 1. If Mary lived in the flat, she must give up the benefits of renting the flat out 2. If Mary rents the flat, she must give up the benefits of living in the flat.*
FOR THIS EXAMPLE ONLY:
Possible benefits of living in the flat:

- *Have a secure place to live*
- *Cook in home* $\longrightarrow$ *more healthy diet, lower input in food*
- *Make friends with neighbors*
- *etc...*

Possible benefits of renting the flat:

- *No need to clean the room*
- *Get extra income by gathering rents* $\longrightarrow$ *more money to fulfill wants*
- *No need to do housework*
- *etc...*

Example 4 ***(Answer)*** *Less income, need to clear the flat, need to do housework.*

When a choice has been made, there must be an opportunity cost.

1.3 Expectation

Expectation affects the satisfaction of the user.

Theorem 1.4 *The higher the expectation of the user, more harder to satisfy the user, higher the accuracy of the answer towards the expectation.*

1.3.1 Applications of Expectation in this algorithm

Since while generating the picture, even the user does not know what is the expectation, any result in a specific range will be accepted. We create a circle with a radius $R = 0.05n^{0.5}$, where n is the dimension. This is the predicted area of acceptable value based on the complexity of the vector the user wants.

1.3.2

1.4 Want

Want is one of the earliest concepts to be discovered in the economy. When people have wants, they need to fulfill their wants by obtaining it. At first, people's wants are food and water. As time goes on, people's want exceeds people's ability to fulfill their wants. That later forms the economy.

Concepts In the modern economy, the concept of want is defined as $Price \div Quantity$.

Theorem 1.5 *Law of demand*

$$Q_d = Q_d(0) - \frac{\partial Q_d(P)}{\partial P} \tag{1.4.0.0.1}$$

where P is Price value, Q_d is Quantity in Demand, a is the Quantity when P=0,

The graph formed by $Theorem 1.1$ is called the "Demand Curve". It shows the quantity of demand over time. Want is affected by several factors. While it could be expressed,

Introduction

1.4.1 Definition of Want

Want is defined as the 'like' or the 'preference' of the model user. This algorithm eventually manages to visualize the tendency of Want through a continuous improvement feedback loop. When the picture generated by this algorithm falls within the user's expectation circle (or meets with the user's expectation), Want is achieved.

1.4.2 Applications of Want in this algorithm

In Figure 4, the vector of change is the want of the user. We use the vectors in each choice to help predict the final target of the user.

1.4.3 Double coincidence of wants

2 Ways to Improve Known Results

The following statements are quoted from (Jiashu Wu, Mathematical Planning and Combination Optimising, 2001):
The fastest way is to use a quick decent algorithm

$$E(x^{k+1}) \leqslant \left\{ 1 - \frac{4aA}{(A+a)^2} \right\} Ex^k = \frac{A-a}{A+a}$$

$^{2}E(x^{k})$(2.0.0.0.1)

This means that when a and A get closer to each other, the velocity of the descent is the greatest. If $\delta_{a,A} \uparrow$, the difference becomes larger, and the velocity is slower.

3 Fundamentals of Stable Diffusion

Stable Diffusion model is an integrated model based on Random Diffusion and Natural Language Processing. The user type in the prompt of what they thought and the model will generate a picture due to the prompt entered.

Example 5 *1. input("Apple eating Banana")*
2. output(a picture of Apple being eaten by Banana)

Example 6 *Input "Ultra man" into SD model*

Part III
Methods

In this section, we will detail the method of how the model works, and what is the principle behind it.

4 Image finder

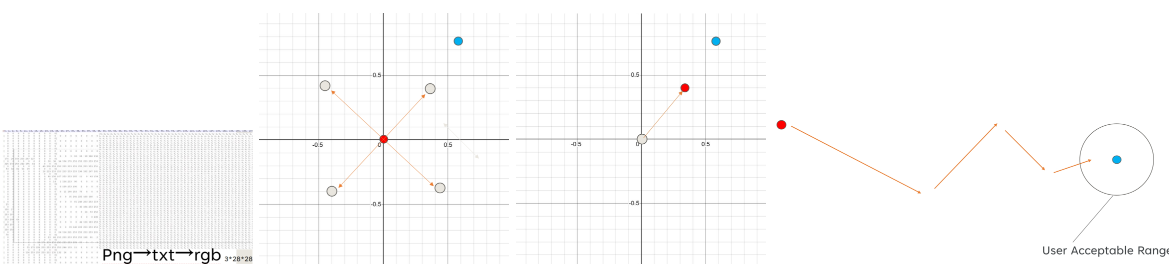
The algorithm will try to guess the preferred image in several ways:

- The user's preference image cannot overpass a limit. Such as all the possible results of the image guessed under an area of a circle starting from the first image guessed.
- A line would pass through the current generated image. Two new images will be generated equal length apart from the first image on the line.
- The user would have to choose one of the images. When one is chosen, the other part of the line, starting from the first image, would be excluded from consideration.

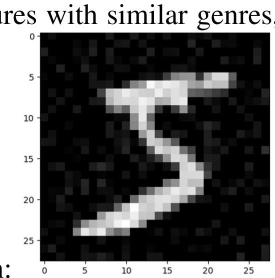
  Definition 4.1 *The preferred image must not be located under the excluded part*

  This is because the user chose the image on the other part of the image, meaning that the preferred image is on the side side of the string together with the chosen image.

- Then, a new line, perpendicular to the line the image is staying on, will generate passing through the selected image. This algorithm is repeated starting from excluding the not preferred part until the selected image arrives at the user's range of preferences or the user manually stops the process.



5 Image generation

For testing purposes, we find some pictures with similar genres, but totally different appearances. For instance, the image used in the introduction:  The vector of the picture is then composed into the image. Noises could be observed in the picture as the remain of "0". The image is composed by refilling the array into an RGB format, then converted into PNG format. In our plan, a prompt should have been generated by directly reading the array generated. The change in vector is described in a signal for LLM to integrate the prompt.

Part IV

Conclusions and Future Works

In conclusion, we created an algorithm to improve the quality of the picture generated by the Stable diffusion model. We are able to excavate users' preferences and hope to improve the algorithm even more. During later investigations, we wish to generate prompts as we planned.

In order to improve the algorithm, currently, two ways have been thought of, which are allowing the user to choose two pictures at once, and to use a decision tree network to enhance the decision prediction.

For future work, we decided to create an API for the algorithm, such that the algorithm could not only improve the results of images but to transform a set of data from one to another. This is because the purpose of our algorithm is to transfer vectors in $(xy)^{3n}$. In theory, generated contents such as texts, and codes written by programmers, could the algorithm be applied on as long as a function is designed to transfer the contents into vectors in n dimension

Part V
Afterwords

6 Words from people related to this paper

Bryan Chen Yueming This is my first time to be in a project with people in universities. The project is rewarding and exciting. I am able to study college-level knowledge and in a college lifestyle. I enjoyed the time with the professors and my classmates. Professor Dr. Haifeng and TA Dr. Junjie (Whom I accidentally met at the City University of Hong Kong during a summer internship) are always there for help and tutors. They taught me a lot of knowledge and concepts containing both theoretical and applied theory. My teammates, Junyi and Dongyi, are very helpful friends. Dongyi has a vast amount of knowledge and thoughts during the critical thinking process and the preparation process. Junyi shows a high level of English. Together, our team has successfully completed this provoking task.

Jeff Wang Junyi This is also my first time to participate in an independent research project. I am extremely grateful to my two teammates for all their hard work and support during our research, especially Yueming. Without their help, I would not have been able to successfully complete our paper on enhancing deep-learning models. We faced many challenges along the way, but through teamwork and determination, we were able to overcome them. This experience has definitely further intrigued my interest in the field of algorithms and machine learning, and I am eager to continue further discussion and collaboration with my teammates.

Dong Yi Completing this paper with my young collaborators has been an incredible journey. As a college student, I was initially unsure about working with younger students, but their enthusiasm and fresh perspectives have truly impressed me.
Throughout the process, I've witnessed their dedication, curiosity, and eagerness to learn. It has been a pleasure guiding them through the complexities of algorithms and witnessing their growth. Their creative ideas and unique insights have also challenged my own thinking and enriched the final outcome of our collaborative effort. It is refreshing to see their hunger for knowledge and their determination to make a meaningful impact on the world.
I am incredibly proud of what we have accomplished together. Our collaboration has produced a comprehensive paper on improving deep-learning algorithms. The experience of working across different age groups and academic levels is definitely something I don't have much chance to experience
To with my two collaborators, thank you for your dedication, hard work, and willingness to explore the unknown. I really look forward to future collaborations that further push the boundaries of our understanding.

7 Special thanks

Special thanks to Professor Prof. Haifeng Xu from the Department of Computer Science & Data Science Institute at the University of Chicago. He guides us through the paper.
Special thanks to Teaching Assistant Dr. Junjie Chen from the Department of Data Science

8 Credits

Credit for Overleaf for LaTeXeditor and BiBTex editor
Special credit to Neoscholar's Cathaypath Institute of Science Program(CIS program). Without this great opportunity, it would be impossible for us to even meet each other.
Special credit to Neoscholar's International Elite Mentorship Program(IEMP). It allows us to enhance our study in Economy and ML.

Acknowledgement

Professor Dr.Haifeng Xu(Professor leading paper)
Teaching Assistant Dr.Junjie Chen(Teaching Assistant of Dr.Haifeng Xu in this project)
Student Chen Yueming(Student in this project)
Student Wang Junyi(Student in this project)
Student Dong Yi(Student in this project)

Reference

References

[1] Lingjiao Chen, Matei Zaharia, and James Zou. 2020. FrugalML: How to Use ML Prediction APIs More Accurately and Cheaply. arXiv:2006.07512 [cs.LG]

[2] Shaddin Dughmi and Haifeng Xu. 2016. Algorithmic Bayesian Persuasion. arXiv:1503.05988 [cs.GT]

[3] Samira Pouyanfar, Saad Sadiq, Yilin Yan, Haiman Tian, Yudong Tao, Maria Presa Reyes, Mei-Ling Shyu, Shu-Ching Chen, and S. Iyengar. 2018. A Survey on Deep Learning: Algorithms, Techniques, and Applications. *Comput. Surveys* 51 (09 2018), 1–36. `https://doi.org/10.1145/3234150`

[4] Yunlong Wang, Shuyuan Shen, and Brian Y Lim. 2023. RePrompt: Automatic Prompt Editing to Refine AI-Generative Art Towards Precise Expressions. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. ACM. `https://doi.org/10.1145/3544548.3581402`

[5] Jiashu Wu, Hao Dai, Yang Wang, Shigen Shen, and Chengzhong Xu. 2022. PECCO: A profit and cost-oriented computation offloading scheme in edge-cloud environment with improved Moth-flame optimization. *Concurrency and Computation: Practice and Experience* 34, 22 (jul 2022). `https://doi.org/10.1002/cpe.7163`

2
I
II
IV
III
[3]
[1]
[2]
[4]
[5]
http://ctan.org/pkg/multirow
https://zhuanlan.zhihu.com/p/583124756
Website for papers