Proposal

Jacek Lewkowicz[1], Yiqing Hu[1,*]

[1]Faculty of Economics Sciences, University of Warsaw, Poland

Corresponding y.hu7@student.uw.edu.pl

1 Title

I have two titles right now, but I don't which one is better.

One is "Exploring User Preferences in the Online Art Market: A Comparative Analysis of Human-Created and AI-Generated Artwork on pixiv.net".

Another one is "Analysis of User Preferences between AI-generated and Human-Created Artwork: A Case Study of pixiv.net".

2 Background

Generative AI models have revolutionized various markets, including the art market, with their transformative capabilities[1]. However, the influence of AI-generated artwork on user preferences within the online art community remains a topic of significant interest. As the prevalence of AI drawing tools, such as Stable-Diffusion (SD) and Midjourney, continues to grow, it becomes crucial to investigate the different topics of AI-generated images and understand user preferences in comparison to human-created artwork among these topics.

3 Literature Review

A. Zarifhonarvar's study examines how ChatGPT affects the labor market[1]. Econometric techniques still make a significant contribution to analyzing cryptocurrencies market in the AI theme[2]. J. Campos, N.R. Ericsson, et al. demonstrated that cyberspace attracts new talent and great business ideas intended to make global art commerce more versatile and efficient, online art market players alongside providers of online art market data and analytics offer interesting avenues for future research in this sector[3]. P. Fortuna and A. Modliński discovered a phenomenon: humans perceive the value of paintings made by AI as lower than the value of those made by humans when the creator of the work is known[4]. Following their research, we will continue to explore user preferences in the online art market.

4 Objectives

The thesis aims to explore the topics influencing the ranking of these images and examine the preferences of ordinary users towards AI-generated versus Human-Created artwork. Figure 1 shows the whole picture of how we tackle the goal.

Through the use of this econometric model such as the ordered probit/logit model, we will quantitatively analyze the rankings of images on both the human-created and AI-generated top lists

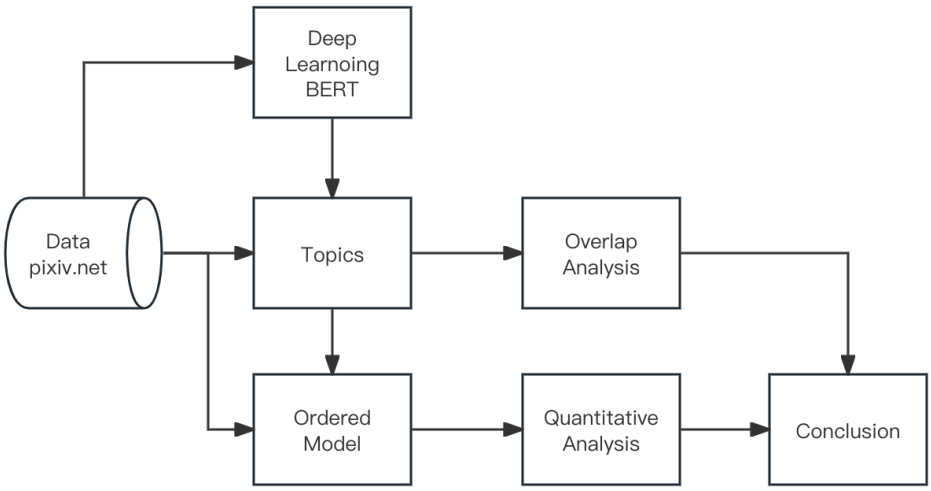

Figure 1: Structure of Project

on pixiv. This analysis aims to identify which topic significantly influences the rankings and the topics are selected by a deep learning model. By examining the correlation between topic and rankings, we are going to investigate whether users have different preferences for human-created and AI-generated artworks. If such different preferences exist, we can leverage these differences to strategically produce and promote human-created and AI-generated artworks in the painting market. However, if the preference differences are minimal, we can assert that human-created and AI-generated paintings are in direct competition. This competition is likely to impact the entire art market.

5 Dataset

In the second semester, I had a web scraping course in which I developed a web spider to crawl the dataset from pixiv.net, a prominent online platform where artists share their artwork. This web spider and the data I got is the final project of this course. The website pixiv.net is widely recognized and highly regarded within the online art community, making it an ideal source for collecting data related to AI-generated artwork and user preferences.

I collected the data spanning from October 31, 2022, to May 15, 2023 from the top list of AI-generated and human-created image pages containing features such as the picture, tags of the pages, dates, statistical data, etc al. After de-duplicating the same image pages which may appear at top with different ranks and different days, we gathered the samples:

- Number of all samples: 14576
- Number of samples of AI-generated Artwork: 8092
- Number of samples of Hand-drawn or human-created Artworks: 6484

6 Methodology

6.1 Machine Learning Algorithms

I will utilize a deep learning model, BERT, to perform topic modeling on mixed data of texts which are tags and images to derive distinct topics between human-created and AI-generated. I will finish this part this semester as my final project in the course of text mining. Figure 2 shows how BERT deals with topic modeling in our project.

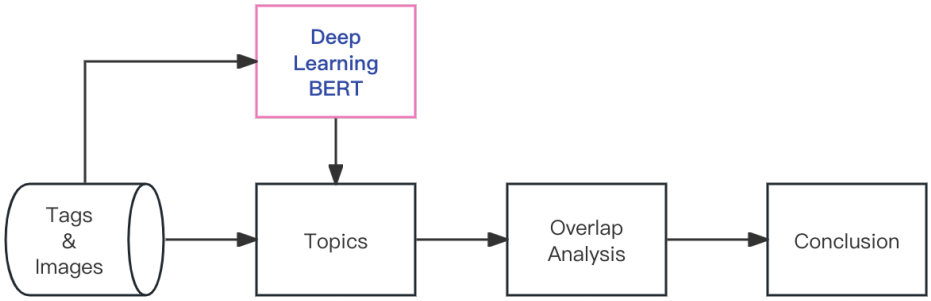

Figure 2: Topic Modeling

Furthermore, if possible, I will examine the extent to which AI and human-created artworks overlap in these topics. Since our data comes from the Top rankings on pixiv, only popular images among users appear in the dataset. Therefore, this overlap also represents the overlap in user preferences. If the overlap is minimal, it suggests that AI and human-created artworks attract different sets of users. Conversely, if there is a significant overlap, it indicates a substantial competitive relationship in user preferences between AI and human-created artworks.

6.2 Econometrics Tools

I will use my outcome of Applied Econometrics' final project to complete this part. In that project, given the nature of the categorized rank as an ordered, continuous, and discrete variable, I employed an ordered choice model in my analysis. Specifically, I utilized both the ordered logit and ordered probit methods to examine the relationships between the independent variables and the categorized rankings effectively. Figure 3 shows the use of ordered models to analyze preference.

In my final thesis, I will add topics from section 6.1 as independent variables into the ordered choice model to find the topics and the categorized rankings that exist in relation significantly. So I have to re-write the Applied Econometrics final project for my thesis.

7 Significance of the Study

This study may provide in-depth insights into the digitization trends in the online painting art market. and explores the role of artificial intelligence in artistic creation, offering not only references for the future development of digital art but also a new and profitable perspective on potential user preferences in the future art market.

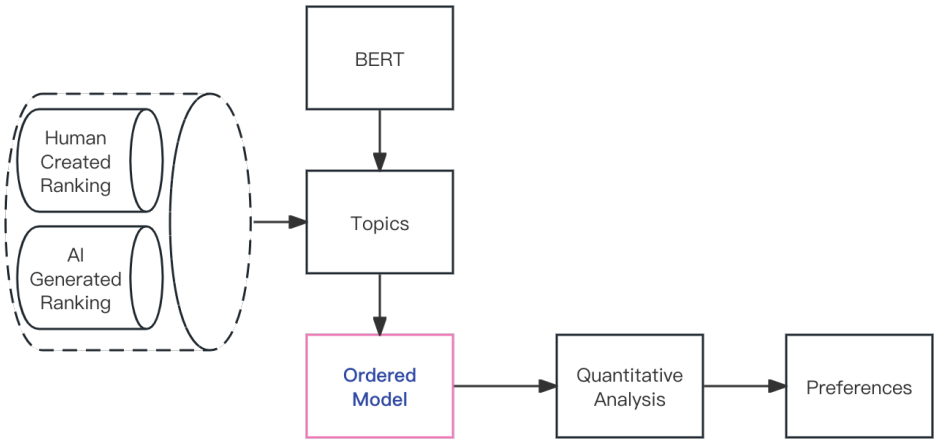

Figure 3: Preference Analysis

8 Expected Results

The accomplishment of this thesis can determine whether art generated by artificial intelligence, especially AI painting is subject to cognitive biases due to technological characteristics. We may also identify the preferences of the general ACG (Animation Comic and Game) users for AI-generated painting works and reveals the potential changes in digital technology on the online art market through this research.

9 Bibliography

[1] A. Zarifhonarvar, "Economics of ChatGPT: A labor market view on the occupational impact of artificial intelligence," *SSRN*, 2023, Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4350925

[2] L. Ante and E. Demir, "The ChatGPT effect on AI-themed cryptocurrencies," *SSRN*, 2023, Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4350557

[3] E. Sidorova, "The cyber turn of the contemporary art market," *Art Markets and Digital Histories*, Jul. 2019, Available: https://doi.org/10.3390/arts8030084

[4] P. Fortuna and A. Modliński, "A(i)rtist or counterfeiter? Artificial intelligence as (d)evaluating factor on the art market," *The Journal of Arts Management, Law, and Society*, vol. 51, pp. 188–201, Mar. 2021, Available: https://doi.org/10.1080/10632921.2021.1887032