# Yu-Gi-Oh! Image Recommendation System

Gorvien Mathis

2023

## 1 Introduction

The goal of this project was to develop an image recommendation system for Yu-Gi-Oh! cards based on user preferences. The project involved several tasks including data acquisition, annotation, user preference profiling, data analysis, and visualization. This report details the methodology and results of each task and provides a self-evaluation of the work, as well as suggestions for future improvements.

## 2 Data Acquisition and Annotation

### 2.1 Data Collection

A dataset containing 8373 Yu-Gi-Oh! card images and associated metadata was sourced from Kaggle, created by Archan Ghosh and Madhurima Maji under the CC0: Public Domain license. The images were downloaded and stored in a folder named "images." Python's 'requests' library was used to automate the process of downloading the images from their source URLs.

### 2.2 Image Metadata

For each image, metadata such as the image format, size, orientation, creation date, card number, card name, card type, attribute, rarity, card set, characteristic, and level were collected and stored in JSON files. Python's 'Pillow' library was employed to extract the metadata directly from the image files, while additional information was gathered from the accompanying CSV file.

### 2.3 Labeling and Annotation

The K-means clustering algorithm was used to identify the dominant colors in each image. The 'scikit-learn' library was utilized for implementing the K-means algorithm, while the 'Pillow' library was used for image processing. Additionally, a CSV file containing further information about the images was utilized.

# 3 User Preference Profiling and Data Analysis

## 3.1 User Preferences

Users were asked to select their favorite images and add tags. Based on their selections, user preference profiles were created, collecting information such as card name, card set, rarity, card type, attribute, characteristic, level, and dominant color. The user input was handled using Python's built-in input functions, and user preference profiles were stored as dictionaries for efficient data retrieval and processing.

## 3.2 Data Analysis

The system analyzed user preferences and their favorite images to provide personalized recommendations. A combination of similarity scoring and collaborative filtering techniques was employed to generate recommendations. The similarity between user preferences and image attributes was calculated using the Jaccard similarity coefficient. Additionally, collaborative filtering was applied to identify users with similar preferences and suggest images that were highly rated by those users.

# 4 Data Visualization and User Interface

## 4.1 Data Visualization

The system visualized various characteristics of the downloaded images, such as the number of images available for each tag. Users could also visualize information related to their favorite images and their user profiles. The 'matplotlib' and 'seaborn' libraries were used to create various types of plots, including bar charts, pie charts, and heatmaps.

## 4.2 User Interface

The user interface was designed to be intuitive and user-friendly, providing functionality for users to visualize information related to their favorite Yu-Gi-Oh! cards and their user profiles. This interactive component helped users better understand the recommendation system and its underlying principles. The interface was built using Python's 'tkinter' library, which allowed for the creation of responsive and visually appealing GUI components.

# 5 Conclusion

The project successfully developed a comprehensive image recommendation system for Yu-Gi-Oh! cards that catered to individual user preferences. Through a combination of user preference profiling, data analysis, and visualization techniques, the system was able to provide personalized recommendations to users.

However, there is room for improvement in the image processing and analysis steps, particularly in refining the K-means algorithm to better focus on the main elements of the card images.

Despite some challenges in understanding the data mining models and finding an appropriate dataset, the project was a valuable learning experience. The practical sessions and exercises aided in building a solid foundation in data mining techniques, which contributed to the success of the project.

Future work on this project could involve implementing advanced image processing techniques to improve the accuracy of the dominant color extraction, as well as exploring other machine learning models for generating more accurate recommendations. Additionally, optimizing the execution time of the project would greatly enhance its usability.

In conclusion, the Yu-Gi-Oh! Image Recommendation System effectively demonstrated the power of data mining and machine learning techniques in providing personalized content recommendations. As the project continues to evolve, it has the potential to become an even more powerful and efficient tool for Yu-Gi-Oh! card enthusiasts.

## Acknowledgements

I would like to express my special thanks to Théotime Perrichet, who has been incredibly helpful and supportive throughout the project.

## Bibliography

## References

[1] Archan Ghosh and Madhurima Maji, *Yu-Gi-Oh! Database*, Kaggle, `https://www.kaggle.com/datasets/archanghosh/yugioh-database`.

[2] Yugipedia, *Yugipedia*, `https://yugipedia.com`.

[3] J. MacQueen, *Some Methods for classification and Analysis of Multivariate Observations*, Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, 1967, pp. 281-297.

[4] Scikit-learn: Machine Learning in Python, Pedregosa et al., Journal of Machine Learning Research, 12, 2011, pp. 2825-2830.