

CSC111 Project Proposal:

Animmend - An interactive anime recommendation system

Ching Chang Letian Cheng Arkaprava Choudhury Hanrui Fan

Monday, March 15, 2021

Problem Description and Research Question

1. Overview of Background

Anime is a genre of film and television animation that originated in Japan. Its wide range of storyline and unique art style has attracted over 20 millions audiences since 2018 (Anime Motivation, 2018), and has been significantly scaling in multiple directions in the past 17 years (Ellis, 2018). With over 210 anime published last year in 2020 (Wikipedia contributors, 2021), there are currently over 17,548 anime in the world (MyAnimeList, 2021) — way too many for anyone to pick!

2. Problem and Motivation

This is a problem for both the audience and the producers of the anime who put in their blood and sweat into creating the work, only for it to be buried down in the monstrous pile of anime that are endlessly being published. Although there are anime-recommendation websites and applications out there that recommend users anime they might like based on their watch history, they tend to favour the trend and popular choices. This causes the established studios to self-perpetuate on the top of the pyramid scheme, while the rest that fails to make a good first impression end up at the bottom of the iceberg, forgotten. Much like every other film and television animation, anime is a form of fine arts medium where the authors, artists, and the production team splash their canvas with imagination and creativity. In a field with such freedom and originality, every anime has a meaningful story for someone in the world. Even the most terribly rated anime by the general could be a treasure to the right audience. However, it is difficult finding this so-called “right audience”. Anime with big names speak for themselves, so what the anime community lacks is a bridge between the unheard anime and the audience.

3. Our Goal

To help mitigate this problem, we seek to create **an application that recommends unpopular anime based on the user’s input anime and relevant themes that they might be interested in.**

Computational Plan

1. Graph Implementation

The recommendation model will be initialized using more than 30,000 anime data taken from the Manami database. Each anime from the Manami database contains one or more “tags” that label the sub-genre, topic, and theme. We will create a function that reads through all 31,166 animes from the Manami database and determines 20 related anime for each anime based on the tags and similarity. Each anime and their related anime, as well as their corresponding tags, will be saved into a **JSON** file to be processed prior to the production stage. This shortens the time needed to construct the graph for users.

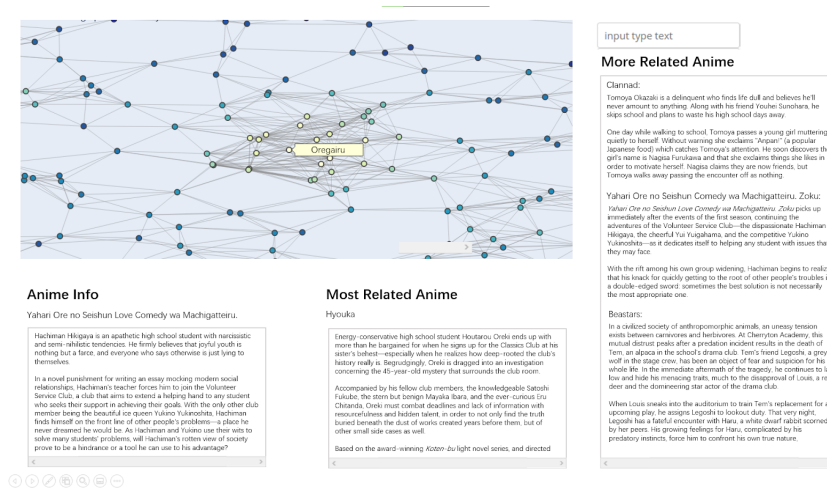
We use a graph structure to represent the collection of anime in our database. Each vertex in the graph represents a single anime. We create a custom class “Anime” that contains a string of an anime’s name, and a list of values representing the relevance of each “tag” for that particular anime (the relevance are ordered in a way such that all they represent the tags in a certain order defined by us). The similarity of the anime with other anime can be seen as the dot product of their ‘tag vectors’ which can be calculated using **NumPy**’s

dot product function, partly inspired by the vector space model used in search engines (Borodin, 2020). The weights are normalized for each anime, and all relevant tags for the anime are weighted equally. As users provide feedback on the recommendations, the program automatically re-adjusts the weighting of each tag using vector normalization, and recomputes the similarity with other anime. An anime’s neighbours are those vertices whose similarities with the dot product greater than a predefined threshold. To simplify the process of visualization, a vertex is restricted to a maximum of twenty neighbours; this implementation keeps users’ convenience in mind, as one would seldom be interested in scrolling past more than twenty recommendations. We can also add new anime to the graph whenever new anime are included in the Manami database; however, as such an operation is $\Theta(n^2)$ (where n is the number of existing vertices), it is not realistic to include in the user interface.

2. Data Example

Data Example saved on GitHub (manami-project, 2021)

3. Visualization and Interactions



Interface Example

There will be an input box where the user can input an anime that they would like to search. Our program would parse the existing **JSON** file that we saved earlier and display all anime connected to the one that the user input, for up to depth 3. The user will be able to click on the nodes of the graph to see an overview of the recommended anime, as well as a ranking of 20 anime ranked by similarity. Each overview of the recommendation comes with a “like” and “dislike” for the user to provide feedback on the quality of our recommendation which will be used to automatically adjust the weighting of the tags, according to the tags that the recommendations correspond to.

4. Technical Requirement

JSON is a format that provides lightweight data interchange originated from **JavaScript** (JSON, 2021). We import the Python library version of **JSON** to exchange and store all the animation data from the Manami database (JSON Python Community, 2021).

Pandas is a Python library written for data manipulation and analysis (Pandas Development Team., 2021). Also, in order to pass data to create visualization. We use **pandas** for **JSON** reading and provide data to graph the relationship between nodes.

NumPy is a Python library that assists scientific calculation (The SciPy community, 2021). In particular, we use **NumPy** to do vector dot products and normalizations, to update similarity (our most important feature) between nodes.

Plotly is a Python open-sourced library that helps to make interactive, publication-quality graphs (Plotly Community, 2021). Specifically, we use **Plotly** to draw the graph and show the relationship between each node by a web page created by **Dash**.

Dash is a Python-based open-sourced library written for building web analytic applications based on **Plotly** (Dash Community, 2021). In particular, we use **Dash** and **Plotly** to build web pages that interact with users

to present new content recommended by our program based on user input and can access our content on any platform.

References

- Anime Motivation. (2018, February 16). *How many people watch anime*. Quora. Retrieved March 13, 2021, from <https://www.quora.com/How-many-people-watch-anime>
- Borodin, A. (2020, October 23). *Great ideas in computing (fall 2020): Week 6*. University of Toronto. Retrieved March 15, 2021, from <http://www.cs.toronto.edu/~bor/196f20/W6.pdf>
- Dash Community. (2021, March 14). *Introduction to dash. plotly*. Plotly. Retrieved March 13, 2021, from <https://dash.plotly.com/introduction>
- Ellis, T. (2018, June 23). *How the anime industry has grown since 2004, according to google trends*. Anime Motivation. Retrieved March 13, 2021, from <https://animemotivation.com/anime-industry-growth-2004-to-2018/>
- JSON. (2021, March 14). *Introducing json*. Author. Retrieved March 13, 2021, from <https://www.cbc.ca/news/world/amazon-deforestation-up-june-1.5644730>
- JSON Python Community. (2021, March 13). *Json encoder and decoder*. Python Software Foundation. Retrieved March 13, 2021, from <https://docs.python.org/3/library/json.html>
- manami-project. (2021, March 11). *anime-offline-database*. GitHub. Retrieved March 13, 2021, from <https://github.com/manami-project/anime-offline-database>
- MyAnimeList. (2021, March 13). *Anime and manga database and community*. MyAnimeList Co. Retrieved March 13, 2021, from <https://myanimelist.net>
- Pandas Development Team. (2021). *pandas documentation*. Pandas. Retrieved March 13, 2021, from <https://pandas.pydata.org/docs/>
- Plotly Community. (2021, March 13). *Plotly python open source graphing library*. Plotly. Retrieved March 13, 2021, from <https://matplotlib.org/>
- The SciPy community. (2021, Jan 31). *What is numpy?* NumPy. Retrieved March 13, 2021, from <https://numpy.org/doc/stable/user/whatisnumpy.html>
- Wikipedia contributors. (2021, March 13). *2020 in anime*. Retrieved March 13, 2021, from https://en.wikipedia.org/wiki/2020_in_anime