Anime Recommender

Alfred Custodio

Bethzaida Guerra

Saatvi R

Thi Quang

Outline

Introduction

Objective

Tableau

Machine Learning

Limitations & Future Work

Q&As

Introduction

What is Anime?

- Anime is Japanese hand-drawn or computer animation. The word is abbreviated pronunciation of “animation” in Japanese, where this term references all animation.

- Outside of Japan, Anime is used to refer specifically to animation from Japan, often characterized by colorful graphics, vibrant characters and fantastical themes.

Introduction- Types of Genre

- Comedy
- Action
- Fantasy
- Adventure
- Sci-Fi
- Drama
- Shounen
- Romance

- School
- Slice of Life
- Supernatural
- Music
- Mecha
- Magic
- Seinen

Introduction- Types of Genre

- Shounen
  - The most popular genre for manga targeting the teenage male audience. Stories usually highlights a youthful hero who possesses a hidden talent that will allow him to surpass himself and accomplish great things.
  - Ex: Dragon Ball, Naruto, One Piece
- Shoujo
  - Similar to Shounen but is targeting female audience. Majority of the stories focus on romantic relationships or emotions.
  - Ex: Sailor Moon, Fruits Basket, Hana Kimi
- Slice of Life
  - Everyday setting that focus on relationships and romantic nature.
  - Ex: Toradora!, K-On!, The Melancholy of Haruhi Suzumiya

- Mecha
  - The word is an abbreviation for “mechanical” in Japanese. Stories encompasses all mechanical objects: cars, guns, computers but majority contains robots.
  - Ex: Gundam, Neon Genesis Evangelion
- Seinen
  - Mainly target male audience between 20-40 years old. Stories contains action, politics, science fiction, fantasy, relationships, sports, or comedy.
  - Ex: Akira, Initial D
  - Josei is the Female version of Seinen

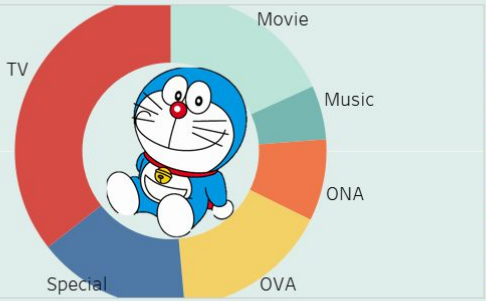
Objective

- Using 2020 data from MyAnimeList, we created a machine learning model that will recommend anime based on the user's input. Recommendation will be drawing from the information provided by the dataset.
- We also utilized Tableau to provide visualization in conjunction with this data.

Tableau

Anime Recommender

User is able to obtain anime recommendation based on the input provided in the "Keyword" section. This is filtered based on the Anime Name, Genre and Synopsis.
User can further break down the recommendation based on the Status of the anime: Complete, Incomplete and Ongoing.



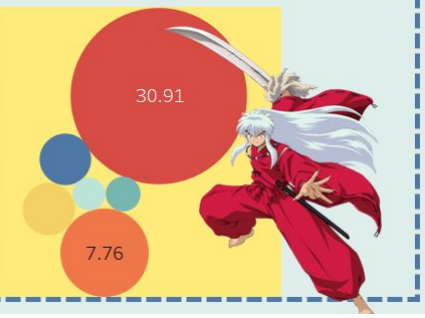
Top Anime per Score

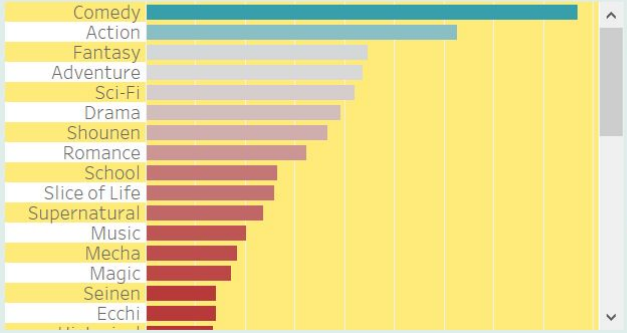
Name	Show	
Fullmetal Al..	True	9.190
Shingeki no ..	True	9.170
Steins;Gate	True	9.110
Gintama°	True	9.100
Hunter x Hu..	True	9.100
Shingeki no ..	True	9.100
Gintama'	True	9.080

Status (All)

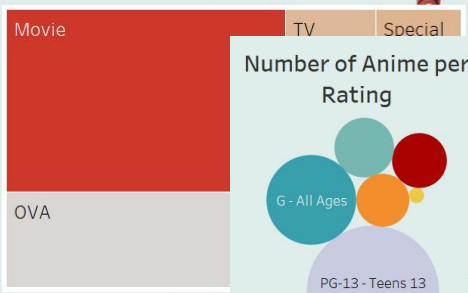
Keyword

Average # of Eps. per Type





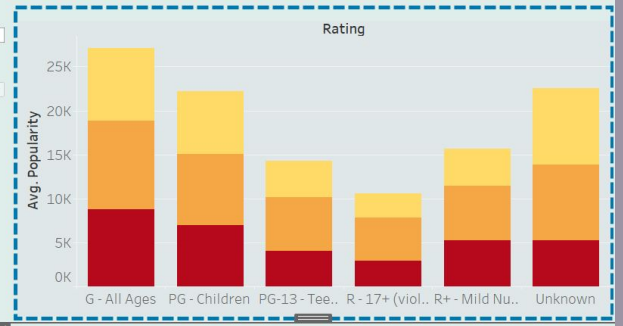
Duration per Episode



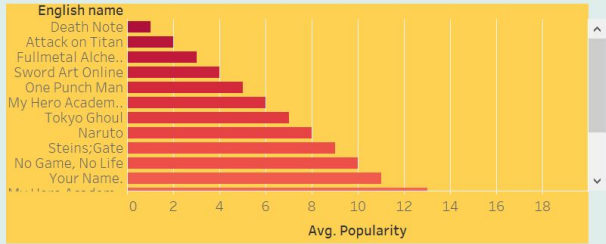
Number of Anime per Rating

Keyword

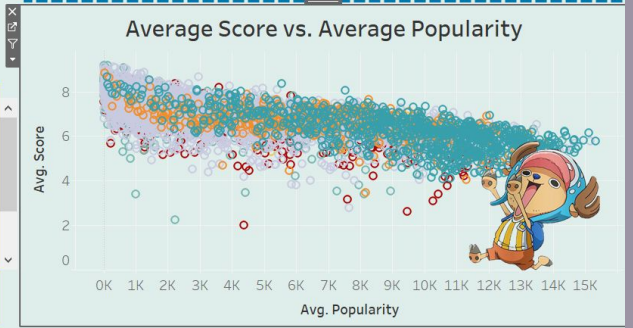
Status (All)



Top 20 Most Popular Anime



Average Score vs. Average Popularity



Machine Learning

Data Preprocessing

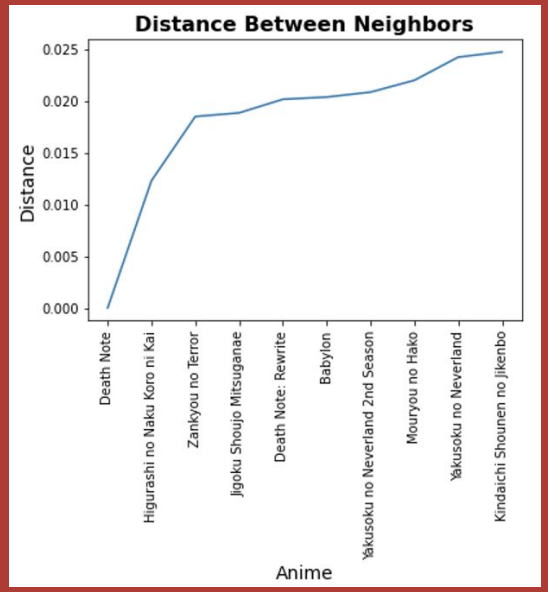
- Review score, maturity rating, and genres were the target features
- Normal one hot encoding was performed on maturity rating
- Manual one hot encoding needed to be done for genres since multiple genres existed for each anime

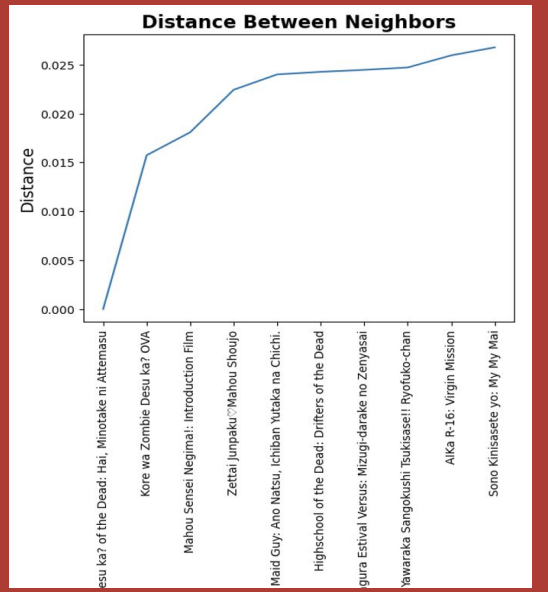
KNN Model (Unsupervised)

- ‘NearestNeighbors’ was used to build the KNN model instead of the usual ‘KNeighborsClassifier’
- Instead of having a train/test split, the entire dataset post-processing was used to fit the model
- Objective was to find the group of anime closest to the one selected in the function

Only Feasible Visualization

- Due to having 50+ dimensions to plot, we could not provide a real representation on how ALL the data points were clustered together
- However, we can observe how close the data points are between the selected and recommended animes





Limitations & Future Work

- Gather up-to-date data using APIs
- Tableau- Search “Keyword” without capitalization/ exact word from the data.
- Make the CSS styling consistent on Recommender, Data Table, and Project Write-Up pages
- Add more inputs to the Recommender page by increasing feature selection and outputting a cleaner table
- Instead of unsupervised KNN, use neural networks to process the huge dimension count and optimize

References

- https://en.wikipedia.org/wiki/Anime
- Data set - https://www.kaggle.com/hernan4444/anime-recommendation-database-2020?select=anime.csv
- KNN Model - https://www.kaggle.com/hernan4444/anime-content-collaborative-knn
- What is KNN - https://realpython.com/knn-python/
- Doing more with KNN - https://scikit-learn.org/stable/modules/neighb
- Choosing best distance metric for Unsupervised KNN - https://www.kdnuggets.com/2020/11/most-popular-distance-metrics-kn

Q&As