ANIME ODYSSEY: EXPLORING GENRES, CLUSTERING & TRENDS

AKSHAT PORWAL (AKSHATP), PRATHAMESH JOSHI (PRATHUJ) & SETTASIT CHITPHENTOM (SETTASIT)

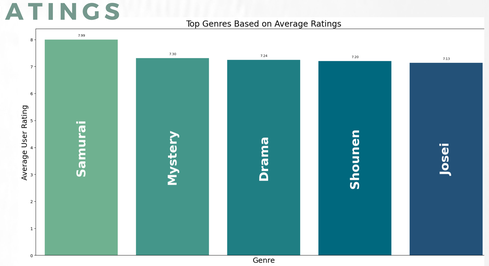
INTRODUCTION

The anime industry has experienced rapid growth in recent years, with an ever-expanding library of anime titles and a growing user base on platforms like MyAnimeList. However, with this growth comes the challenge of providing users with personalized recommendations, enhancing user satisfaction, and understanding the genre associations.

RESEARCH QUESTIONS

1. Which anime genres are most highly rated by users based on average scores?
2. How does the hierarchical clustering of anime genres and user grouping into genre clusters provide insights into diverse user preferences?
3. How does the predictive performance of the recommendation system in rating predictions, and what are the effective strategies to recommend unwatched anime based on user reviews?
4. What is the temporal trend in user preferences concerning the release timing of anime, and to what extent do genres influence user ratings of anime?
5. How do word clouds, derived from anime synopses within specific genres, illuminate the unique themes characterizing each genre?

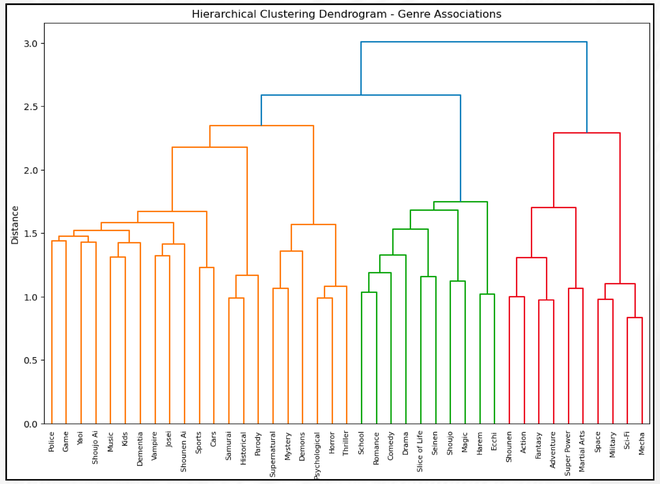
TOP GENRES BASED ON AVERAGE RATINGS

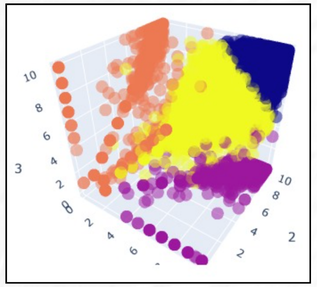


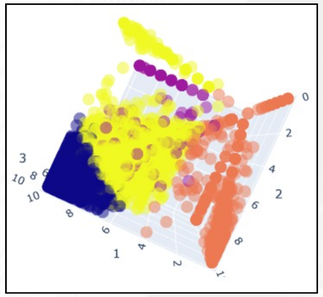
EXPLORING THE JACCARD SIMILARITIES BASED ON GENRES

Anime: Fullmetal Alchemist: Brotherhood	
One Piece: Heart of Gold	0.75
One Piece Movie 14: Stampede	0.66
Akame ga Kill!	0.62

ANALYSING USER CLUSTERS



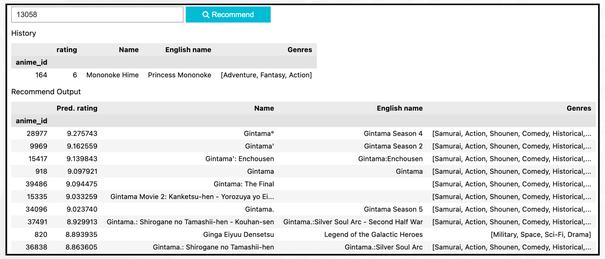




USER RECOMMENDATION

Prediction performance is evaluated via 3-fold cross-validation on reviews from ~317,000 users (13.4 million reviews). After training on the complete dataset, the model can predict ratings for unwatched anime, facilitating personalized recommendations based on anticipated scores.

Recommendation Model	Mean Absolute Error (MAE)
Singular Value Decomposition (SVD)	0.88
K-Nearest Neighbors	0.95



13058 | Recommend

History

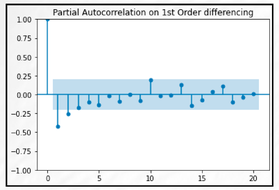
anime_id	rating	Name	English name	Genres
164	6	Mononoke Hime	Princess Mononoke	[Adventure, Fantasy, Action]

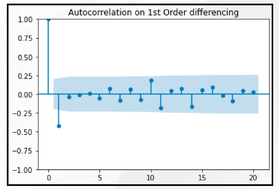
Recommend Output

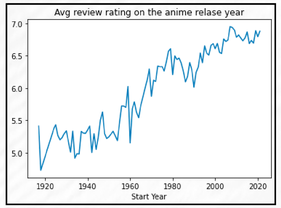
anime_id	Pred. rating	Name	English name	Genres
28977	9.275743	Gintama°	Gintama Season 4	[Samurai, Action, Shounen, Comedy, Historical,...
9969	9.162559	Gintama'	Gintama Season 2	[Samurai, Action, Shounen, Comedy, Historical,...
15417	9.139843	Gintama°: Enchousen	Gintama:Enchousen	[Samurai, Action, Shounen, Comedy, Historical,...
918	9.097921	Gintama	Gintama	[Samurai, Action, Shounen, Comedy, Historical,...
39486	9.094475	Gintama: The Final		[Samurai, Action, Shounen, Comedy, Historical,...
15335	9.033259	Gintama Movie 2: Kanketsu-hen - Yorozuya yo Ei...		[Samurai, Action, Shounen, Comedy, Historical,...
34096	9.023740	Gintama.	Gintama Season 5	[Samurai, Action, Shounen, Comedy, Historical,...
37491	8.929913	Gintama.: Shirogane no Tamashii-hen - Kouhan-sen	Gintama.:Silver Soul Arc - Second Half War	[Samurai, Action, Shounen, Comedy, Historical,...
820	8.893935	Ginga Eiyuu Densetsu	Legend of the Galactic Heroes	[Military, Space, Sci-Fi, Drama]
36838	8.863605	Gintama.: Shirogane no Tamashii-hen	Gintama.:Silver Soul Arc	[Samurai, Action, Shounen, Comedy, Historical,...

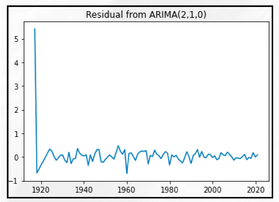
USER TEMPORAL PREFERENCE

Differencing of the average ratings on release year results in a weakly stationary time series. Investigation of partial autocorrelation in the differenced data indicates that an ARIMA(2, 1, 0) model is well-suited for capturing the underlying patterns in the pre-differenced dataset, showcasing its effectiveness in modeling the temporal dynamics of the data.

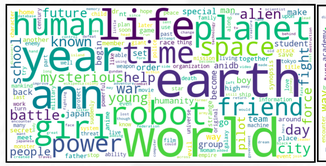




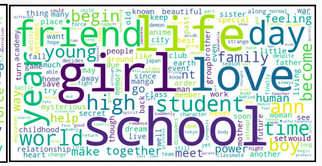




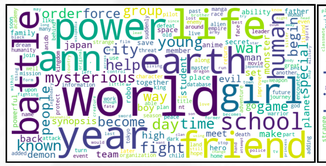
GENRE- SPECIFIC WORD CLOUDS



Genre: Sci-fi



Genre: Romance



Genre: Action

Genre: Comedy

FINDINGS

- User can be classified into 4 groups based on how users rate the anime.
- There are multiple approaches in recommending anime to users; Using similarity of genres, or recommendation system which based on peer reviews.
- The average rating based on release year of anime fits ARIMA(2,1,0) can be used to predict average rating of future anime.
- Word Clouds suggest genres associated with certain keywords. Yet, there are keywords which universally associated with anime itself.