Jason M. White
ECE 479
Final Project

Anime Rankings Using Aspect Based Sentiment Analysis

Introduction

This project explores whether anime rankings can be improved using Aspect-Based Sentiment Analysis (ABSA), fuzzy logic, and Word2Vec. Review data for five anime titles were scraped from MyAnimeList (MAL), and aspect terms were extracted and classified for sentiment using ATE and ATSC. Word2Vec was used to group similar terms into five canonical categories—Narrative, Visual, Audio, Character, and Miscellaneous—which were then weighted and combined using fuzzy logic to produce an overall rating. Compared to MAL's official scores, the ABSA-based ratings more accurately reflected the sentiments expressed in user reviews, offering a more nuanced, data-driven approach to anime evaluation.

AI and ML Concepts Applied

This project uses three key concepts: Aspect-Based Sentiment Analysis (ABSA), Word2Vec, and fuzzy logic. ABSA includes two main steps: Aspect Term Extraction (ATE), which identifies specific features mentioned in a review (e.g., "story", "music"), and Aspect Term Sentiment Classification (ATSC), which determines the sentiment (positive, neutral, or negative) for each aspect. Sentiments are then converted to scores, positive = 1, neutral = 0.5, negative = 0, and grouped by related themes.

To group similar aspect terms, Word2Vec maps words into a vector space where semantically similar words are closer together. For example, "story" and "narrative" will appear near each other. Word2Vec is used to match each aspect term to one of five predefined categories: Narrative, Visual, Audio, Character, or Miscellaneous.

Fuzzy logic is then applied to assign each aspect term to the category it most closely resembles based on Word2Vec similarity scores. For example, if "art" is most similar to "Visual," its sentiment score contributes to that category. This process is repeated across all reviews to calculate weighted category scores and produce the final rating.

Methodology

This project performs Aspect-Based Sentiment Analysis (ABSA) on reviews scraped from MyAnimeList (MAL) for five anime titles: Naruto: Shippuden, Tokyo Ghoul, Yuru Camp, Kaguya-sama: Love is War, and Mushoku Tensei. Each title was chosen for having at least 14 pages of diverse user reviews.

I used a web scraper to collect reviews, then cleaned them by extracting the user's recommendation label and the full text (including hidden "read more" content).

Using the InstructABSA model (for ATE and ATSC subtasks), I performed sentiment analysis to extract aspect terms and classify their sentiment. Reviews are split into 512-character chunks to fit model input limits. The output for each review included: The predicated average sentiment score, the aspect-sentiment pairs, the original user recommendation, and the review text.
Reviews with no detected aspects were stored separately. Those with valid results were saved to JSON. Next, we:

1. Converted sentiment words to numerical values (positive = 1, neutral/conflict = 0.5, negative = 0).
2. Aggregated all aspect terms and computed their average sentiment.
3. Filtered and sorted these aspects by frequency and average sentiment.
4. Matched each aspect to a canonical category (Narrative, Visual, Audio, Character, Miscellaneous) using either exact string match or Word2Vec similarity via Gensim's KeyedVectors.

5. Calculated an average sentiment for each category.

Finally, I computed an overall anime rating by applying the following weights to each category: Narrative: 25%, Visual: 25%, Character: 20%, Audio: 15%, and Miscellaneous: 15%. The final rating is a weighted average scaled to a percentage.

Project Pipeline and Architecture

The system begins with the extraction of user-generated anime reviews scraped from MyAnimeList (MAL), then cleaned and preprocessed for analysis. The first processing step is Aspect Term Extraction (ATE), which identifies key discussion points within the reviews, such as "story", "animation", or "character design." Each extracted aspect is then passed through Aspect Term Sentiment Classification (ATSC), which determines the sentiment polarity (positive, neutral, or negative) for each aspect in context of its sentence. These polarities are then mapped to numerical values: 1 for positive, 0.5 for neutral, and 0 for negative.

To reduce noise and semantically group similar terms, Word2Vec is used to calculate vector similarities between each extracted term and a set of five canonical aspects or a predefined list used for semantic mapping: Narrative, Visual, Audio, Miscellaneous, and Character.

Each aspect is assigned to the canonical category with the highest similarity score. A fuzzy logic approach is then used to assign aspect sentiment scores in cases of ambiguity, selecting the most semantically appropriate match. Once all aspect terms are grouped and scored, each anime receives five aggregated scores—one for each canonical aspect. These are multiplied by pre-defined weights and summed to produce a final ABSA rating.

The results are visualized through four figures to aid interpretation:

1. **Bar Plot of Canonical Aspect Sentiment & Occurrences**. This graph shows the average sentiment for each canonical aspect along with a visual cue for how often those aspects occurred in the dataset.
2. **Bubble Plot of Fine-Grained Aspects**. Each circle represents an individual aspect term, sized by how frequently it occurred and colored by its canonical aspect. This graph reveals the distribution of sentiment values across all fine-grained terms.
3. **Bar Chart: Average Sentiment per Aspect per Anime**. This plot compares how each anime performs across the five canonical aspects, using sentiment percentages for easier comparison.
4. **Line Graph: ABSA vs. MAL Ratings**. This final graph compares the ABSA-derived scores against official MyAnimeList scores, highlighting how user sentiment per aspect differs from overall platform rankings.

Results

Reference (1) Bar Plot of Canonical Aspect Sentiment & Occurrences, (2) Bubble Plot of Fine-Grained Aspects on last page

Naruto Shippuden: A large number of reviewers commented on the Narrative aspect, evidenced by the high occurrence count. Although Naruto is a very popular anime, its Narrative sentiment score is lower than its Audio score. This can be attributed to widespread criticism of filler episodes. In contrast, the soundtrack received strong praise, which is reflected in the high sentiment score for Audio. The fine-grained bubble plot confirms this—large bubbles under Narrative show that users discussed this aspect frequently, often with mixed feelings.

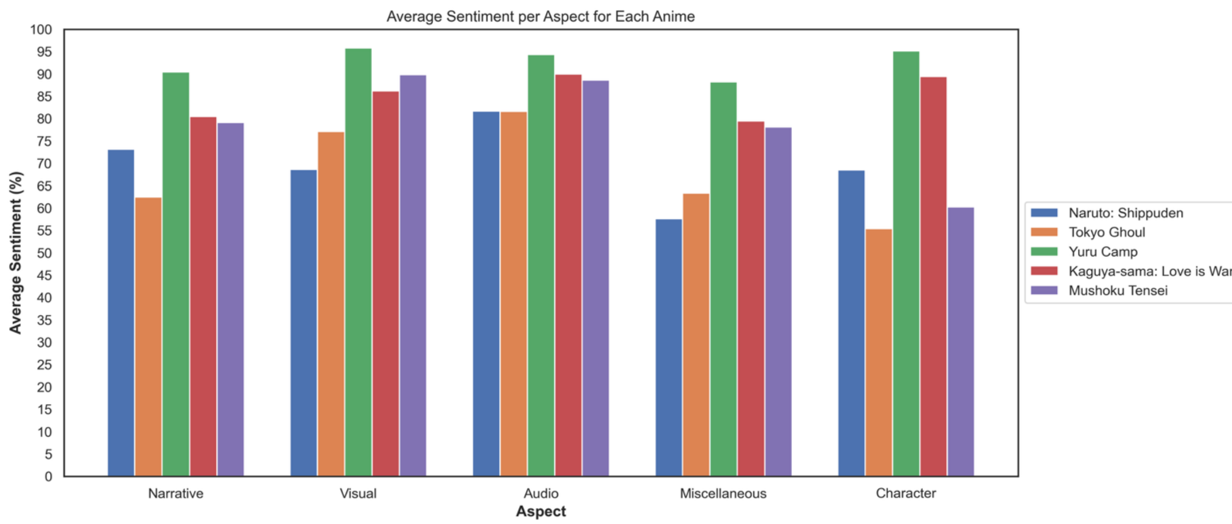
Tokyo Ghoul: Character and Narrative received notably low sentiment scores, while Audio remained high. These results align with critical responses to the anime's divergence from the manga, leading to disappointment in story and character development. However, the soundtrack was widely praised, and this is clearly reflected in the Audio score. Large bubbles in the Narrative section of the fine-grained graph suggest that many users had strong opinions on story quality—even if they were negative.

Yuru Camp: This anime received consistently high sentiment across all canonical aspects, with Character being the only area with fewer mentions. Even infrequent, less common aspects were rated highly in the fine-grained bubble plot. The lack of negative sentiment and near-universal praise suggest this anime is broadly enjoyed by audiences and excels in delivering a consistent, well-received experience.

Kaguya-Sama: Love is War: Although Character scores are high, it was mentioned less frequently than other aspects. Narrative remained a strong focus, receiving high sentiment but also generating diverse opinions. This likely stems from the anime's clever and dialogue-heavy storytelling, which invites different interpretations. The fine-grained plot confirms this with large bubbles under Narrative, showing that this category sparked the most discussion. Both Visual and Audio aspects were well received.

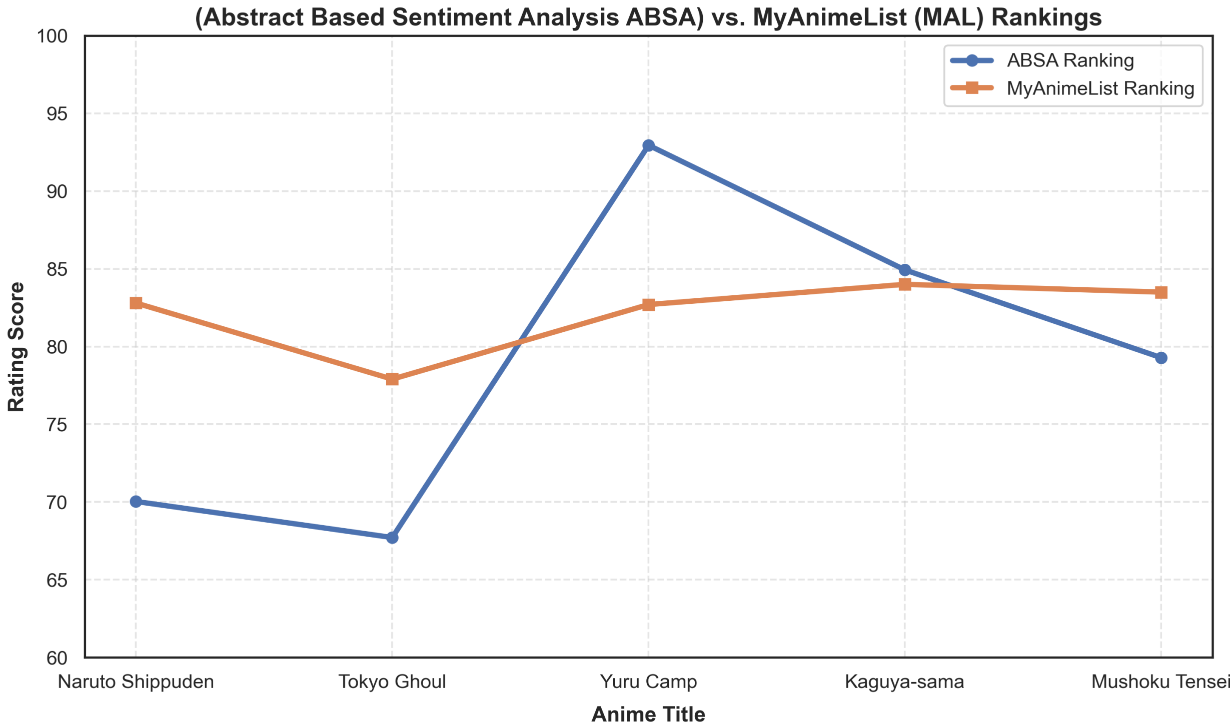
Mushoku Tensei: This anime presents an interesting mix—while Visual and Audio aspects received top scores, Narrative and especially Character are more divisive. The strong reactions to morally complex characters are reflected in the low sentiment for Character. Despite controversy, many praised the animation and soundtrack, leading to dominant Visual and Audio scores. The fine-grained plot shows large bubbles under Visual and Narrative, indicating widespread engagement and strong opinions—whether positive or negative.

Bar Chart: Average Sentiment per Aspect per Anime



This bar chart compares how each of the five anime is rated across the five canonical aspects: Narrative, Visual, Audio, Miscellaneous, and Character. Audio is consistently rated the highest, especially for Yuru Camp, Kaguya-Sama, and Mushoku Tensei, showing strong approval for sound and music. Visual scores are also high for most titles, reflecting positive views on animation quality. Narrative and Character show more variation. For example, Tokyo Ghoul scores low in these areas, likely due to criticism of its story and character development. In contrast, Kaguya-Sama and Yuru Camp are rated highly across most aspects. The Miscellaneous category includes elements like pacing, production quality, and studio reputation. These scores vary, showing that viewers notice and care about these factors, even if they're not always the focus. Overall, the chart gives a clearer picture of what each anime does well and where it may fall short.

Line Graph: ABSA vs. MAL Ratings.



The line graph compares ABSA sentiment-based scores with MyAnimeList (MAL) ratings, revealing meaningful differences. While MAL scores reflect overall popularity, ABSA ratings vary based on sentiment across aspects like Narrative, Audio, and Character. For example, Yuru Camp scores much higher in ABSA (92.95 vs. 82.7), while Naruto Shippuden and Tokyo Ghoul score lower (70.3 vs. 82.8 and 67.7 vs. 77.9, respectively). Kaguya-sama: Love is War is closely aligned (84.9 vs. 84.0), and Mushoku Tensei is slightly lower in ABSA (79.3 vs. 83.5). These differences suggest ABSA captures more detailed feedback from viewers, highlighting strengths and weaknesses that MAL's general scores may overlook.

Conclusion

ABSA ratings offer a more accurate reflection of how viewers truly feel about an anime by analyzing sentiment across its specific canonical aspect categories. Unlike a general 0–10 MAL score, ABSA uses ATE and ATSC to extract and classify real opinions from reviews, revealing deeper insight into what audiences value most.

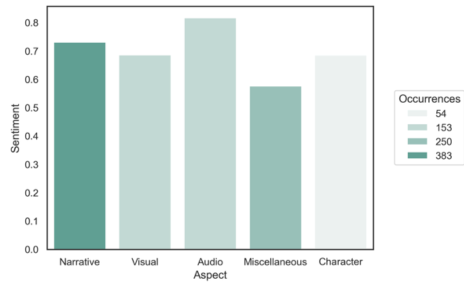
References

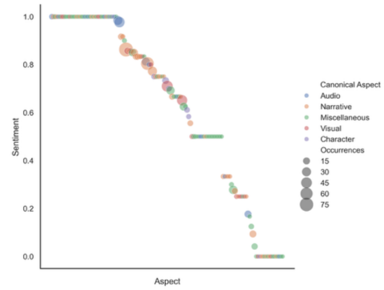
[1] InstructABSA GitHub Repository: https://github.com/kevinscaria/InstructABSA
[2] Word2Vec Archive: https://code.google.com/archive/p/word2vec/
[3] Gensim Word2Vec: from gensim.models import KeyedVectors
[4] Scaria, Kevin, et al. (2023). InstructABSA: Instruction Learning for Aspect-Based Sentiment Analysis. Arizona State University.

Project GitHub: https://github.com/JasonmWhite520/anime_absa_ece479

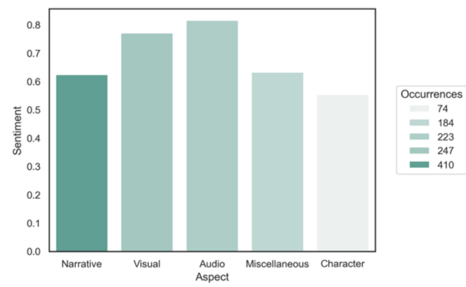
Graphs: (1) Bar Plot of Canonical Aspect Sentiment & Occurrences, (2) Bubble Plot of Fine-Grained Aspects

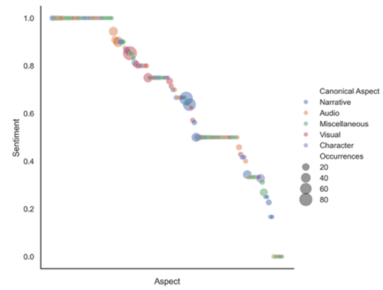
<u>Naruto</u> (1) (2)



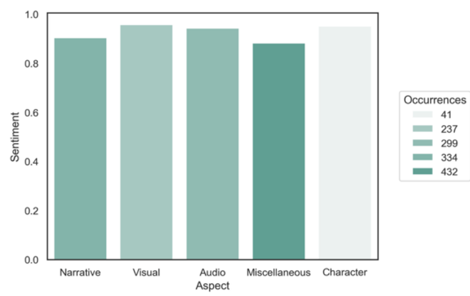


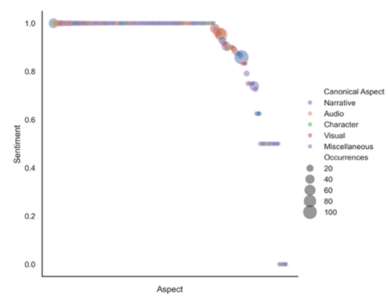
<u>Tokyo Ghoul</u> (1) (2)



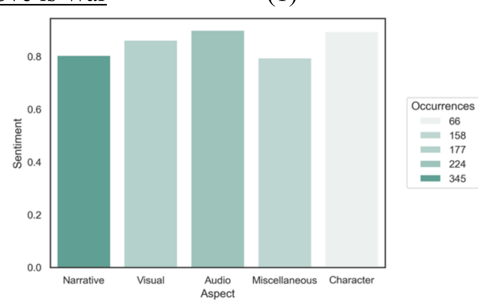


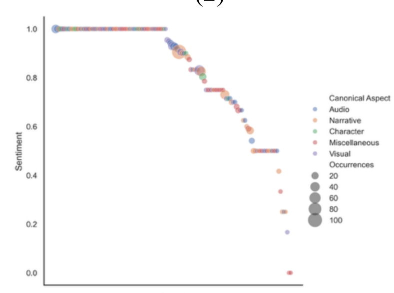
<u>Yuru Camp</u> (1) (2)





<u>Kaguya-Sama Love is War</u> (1) (2)





<u>Mushoku Tensi</u> (1) (2)