MANIPAL INSTITUTE OF TECHNOLOGY
BENGALURU
(A constituent unit of MAHE, Manipal)

Mini Project Report

on

AniSense: An Anime and Manga Recommendation System

SUBMITTED BY

School of Computer Engineering

October 2025

Name: Vishal Agarwal

Reg No: 235890336

1. ABSTRACT

AniSense is a hybrid content-based recommendation system developed to address the growing demand for personalized content discovery in the rapidly expanding anime and manga industry. Leveraging the AniList GraphQL API, the system retrieves real-time metadata for 8,000 titles (4,000 anime and 4,000 manga), ensuring coverage of both popular and ongoing releases.

To capture nuanced relationships between titles, the system employs a multi-aspect similarity fusion model that integrates semantic embeddings, TF-IDF lexical vectors, normalized numeric features, and one-hot encoded categorical attributes into a unified similarity score. A recency-aware boosting mechanism prioritizes contemporary works, while genre-aware re-ranking, fuzzy alias resolution, and manual alias substitution further enhance thematic fidelity and query robustness.

The recommendation pipeline involves comprehensive preprocessing, feature engineering, and similarity computation, deployed through a unified Streamlit application that manages both frontend interaction and backend logic with persistent caching for efficiency. Experimental evaluation demonstrates that AniSense consistently delivers genre-faithful, semantically aligned recommendations with sub-second retrieval latency.

Overall, the system highlights the effectiveness of real-time API integration and hybrid semantic modelling in dynamic content domains, offering a scalable solution for personalized anime and manga discovery.

2. INTRODUCTION

2.1 Industry Background

The global anime and manga industries have witnessed unprecedented growth in recent years. The anime market was valued at over $28 billion in 2022 and is projected to exceed $60 billion by 2030[1]. Similarly, the manga market reached $13.7 billion in 2023 and is expected to grow to $42.2 billion by 2030 [1].

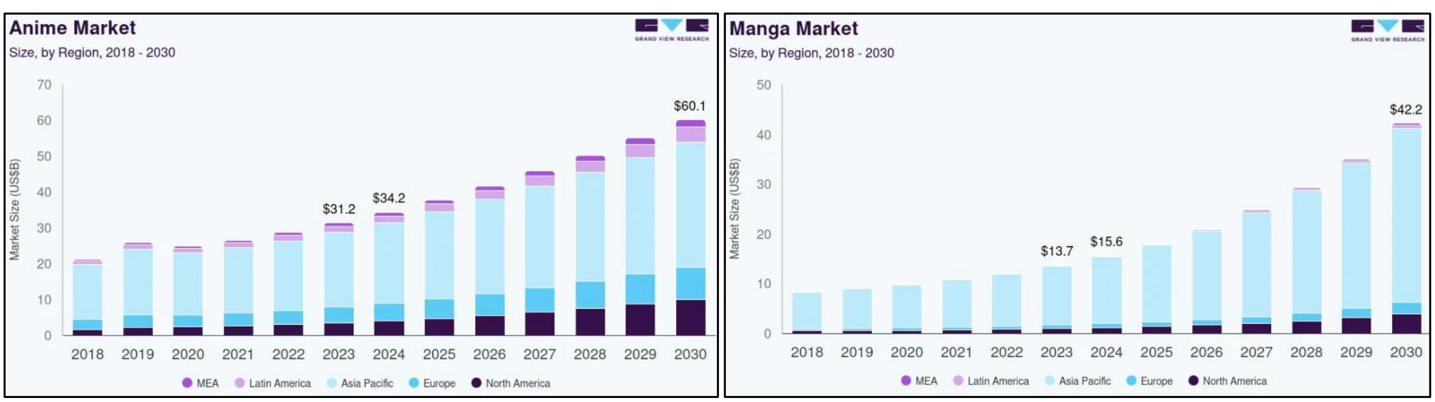

Figure 1: Projected global market size of anime and manga by region (2018–2030), highlighting growth across Asia Pacific, North America, and Europe

Anime has firmly entered the global mainstream, exemplified by the release of Demon Slayer: Infinity Castle Part 1 (2025), which grossed over $600 million worldwide, surpassing its predecessor Mugen Train to become the highest-grossing Japanese film of all time [2]. This success contributed to a resurgence in manga sales, with the franchise exceeding 220 million copies in circulation despite concluding in 2020 which ranks them as the seventh best-selling manga of all time [3]. This dual expansion reflects the interdependent nature of the two mediums, where manga often serves as source material for anime adaptations, creating a continuous cycle of content production and consumption.

Date	Event
Apr 2019	Anime Season 1 airs
Oct 2020	Mugen train movie release
Dec 2020	Manga Ends (Vol. 23)
Dec 2021	Season 2: Entertainment District Arc
Jul 2025	Infinity Castle: Part 1 movie release

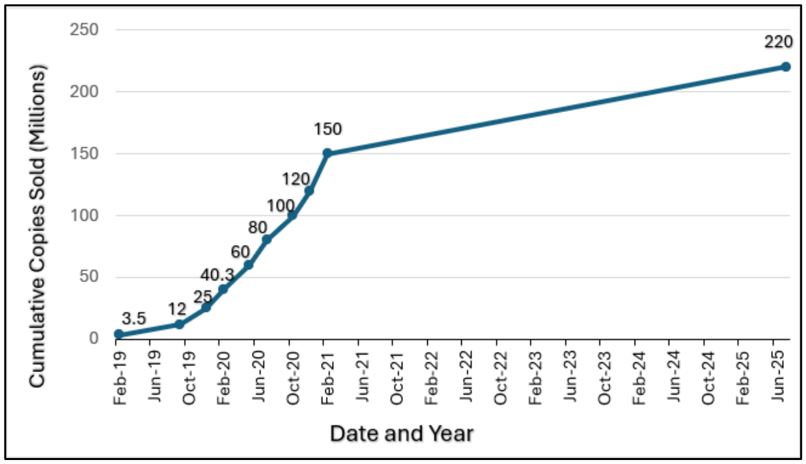

Figure 2: Key events and release timeline influencing cumulative sales of Demon Slayer manga (2019–2025) [4]

Japan's domestic manga market alone generated approximately 677 billion yen (approx. $4.5 billion) in 2022 [5], with digital manga sales accounting for nearly two-thirds of revenue— a clear indicator of shifting consumption patterns [5]. Internationally, regions such as North America and Europe have emerged as significant growth zones, driven by streaming platforms, digital publishing, and global fan communities [6].

This rapid expansion has also created an overwhelming volume of content, making personalized discovery increasingly challenging for consumers. As the industry continues to globalize, the need for intelligent, scalable recommendation systems becomes critical to help users navigate and engage with relevant titles across diverse genres, formats, and languages.

2.2 Related Work

Recommendation systems have been widely studied across domains such as e-commerce, music, and video streaming. In the context of anime and manga, existing systems often rely on collaborative filtering or user rating-based models [7], which require large volumes of user interaction data and suffer from cold-start limitations. Platforms like MyAnimeList and AniDB offer limited recommendation capabilities, typically rule-based or popularity-driven, but lack semantic personalization.

Recent research has explored hybrid approaches which combine content-based filtering with deep learning or graph-based techniques [9]. However, many implementations depend on static datasets, such as the widely used Anime Recommendation database [8] sourced from Kaggle, which lack real-time updates, multilingual metadata, and comprehensive genre tagging. These limitations reduce their effectiveness in dynamic content domains like seasonal anime and ongoing manga series. AniSense addresses these gaps by leveraging the AniList GraphQL API for real-time metadata access and implementing a hybrid similarity fusion model that integrates semantic embeddings, lexical analysis, numeric and categorical features, and genre-aware re-ranking. Unlike others, AniSense delivers personalized, multilingual, and thematically consistent recommendations without relying on user ratings or static datasets, offering a robust and scalable solution for dynamic content discovery.

System/Dataset	Approach	Limitations	AniSense Advantage
MyAnimeList/AniDB	Rule-based/ Popularity	No semantic modelling, cold-start issues	Semantic fusion, no user ratings needed
Kaggle Anime DB	Static content-based	Outdated, limited metadata, no multilingual support	Real-time API, rich metadata
Hybrid DL/ Graph models	Deep learning + CB	Often rely on static datasets, limited alias handling	Real-time, multilingual, alias-aware

Table 1: Comparison of existing anime/ manga recommendation systems and AniSense

2.3 Benefits of AniList API over Traditional Datasets

The decision to use AniList API provides significant advantages:

- Real-Time Data Access
  AniList provides continuously updated metadata [10] for seasonal anime and ongoing manga, ensuring that the recommendations reflect the latest releases and trends. This dynamic coverage is critical for relevance in fast-evolving content domains.

- Rich metadata coverage
  Beyond basic titles and genres, AniList offers multilingual naming conventions (English, Romaji, Native), detailed descriptions, genre and tag annotations, popularity metrics, staff and studio information, media assets such as cover images, banner art, YouTube trailer thumbnails, and additional metadata fields that can be selectively retrieved as needed [10]. This depth supports robust feature engineering and semantic similarity modelling.

- GraphQL flexibility
  AniList's GraphQL interface allows developers to define the exact structure of the data they need [11]. This prevents both over-fetching and under-fetching, resulting in faster, more efficient queries compared to traditional REST APIs.

- Technical Efficiency
  The API's interface enables selective, efficient data retrieval with minimal overhead. Its consistent schema eliminates the need for manual data cleaning or normalization, streamlining preprocessing and improving scalability for large-scale recommendation pipelines.

- Reliability and Documentation
  As an official API, AniList offers stable uptime, version control, and comprehensive documentation [12], making it a dependable foundation for production-grade systems.

- Community-Driven Accuracy
  AniList's database is maintained and actively curated by a global community of contributors, ensuring that the metadata remains accurate, extensive and up-to-date [13]. This collaborative model also improves alias handling and multilingual coverage, which are often weak points in static datasets.

Overall, these advantages make AniList a superior alternative to static datasets. Its real-time updates, rich metadata, technical efficiency, and community-driven data provides the foundation that AniSense requires to deliver dynamic and semantically faithful recommendations across both anime and manga.

2.4 Problem Statement

Anime and manga enthusiasts frequently struggle to discover new content aligned with their preferences due to several key challenges:

- Information Overload
  The sheer volume of available titles across genres, formats, and seasons makes manual exploration overwhelming and inefficient.

- Lack of Nuanced Feature Matching
  Existing platforms often rely on basic genre tags or popularity metrics, overlooking deeper thematic elements, tone, and narrative style.

- Cross-Medium Navigation Barriers
  Users interested in both anime and manga face difficulty transitioning between adaptations, sequels, and spin-offs due to fragmented metadata and inconsistent linking.

- Static and Incomplete Datasets
  Many systems rely on outdated or manually curated datasets, lacking real-time updates, multilingual metadata, and comprehensive tagging
  .

- Limited Personalization
  Most recommendation engines offer generic suggestions based on popularity or user ratings, with minimal semantic understanding.

- Query Robustness Issues
  Informal queries, aliases, and multilingual variations are often not recognized, leading to poor retrieval and mismatched recommendations.

These limitations highlight the need for a scalable, content-aware recommendation system that integrates real-time metadata, semantic modeling, and robust query handling to deliver personalized, genre-faithful suggestions across both anime and manga domains.

2.5 Objectives

- Develop a hybrid content-based recommendation system for anime and manga using real-time metadata from the AniList GraphQL API.

- Integrate multiple feature modalities including semantic embeddings (all-mpnet-base-v2), TF-IDF analysis, categorical encoding, and numeric feature correlation into a unified similarity model

- Implement genre-aware re-ranking and tag-based filtering to enhance thematic fidelity and ensure relevance in recommendations

- Enable robust query handling through fuzzy alias matching and manual alias substitution, supporting informal, multilingual queries and few abbreviation- based user inputs.

- Develop a unified Streamlit application that seamlessly integrates both frontend interaction with backend recommendation logic for real-time user engagement.

- Optimize system performance through persistent caching, modular filtering by media type, and efficient query execution strategies.

2.6 Expected Outcomes

AniSense is expected to deliver personalized, genre-faithful recommendations across both anime and manga domains by leveraging real-time metadata and multi-modal similarity fusion. The system aims to improve user satisfaction through semantic matching, multilingual query support, and robust alias handling. By integrating cross-medium navigation and thematic re-ranking, AniSense facilitates seamless exploration of related titles, including adaptations and sequels. Its modular architecture and caching strategies are designed to ensure scalable performance and efficient query execution, making it suitable for dynamic content environments.

3. METHODOLOGY

3.1 Multi-Modal Similarity Fusion Architecture

The recommendation engine employs a hybrid similarity model that integrates four distinct similarity matrices — semantic, lexical, numeric, and categorical — each capturing a different aspect of content relevance. These matrices are fused using weighted aggregation to produce a unified similarity score that drives recommendation ranking.

Semantic Similarity (40% Weight)

- Utilizes the ‘sentence-transformers/all-mpnet-base-v2’ [14] model to encode combined textual metadata into 768-dimensional embeddings
- Embeddings are normalized and compared using cosine similarity to capture deep semantic relationships between anime and manga titles

Lexical Similarity (5% Weight)

- Applies TF-IDF vectorization with a vocabulary size capped at 5,000 features [15].
- Text inputs are constructed using a weighted combination of key metadata fields: description (25%), genres (35%), tags (20%), studio (3%), source (2%), relations (15%)
- Cosine-similarity is computed over the resulting sparse term-document matrix to capture surface-level textual overlap

Numeric Similarity (10% Weight)

- Involves Min-Max scaling of quantitative features [16] such as meanScore, averageScore, popularity, favourites, duration, episodes, chapters, volumes
- Similarity is computed using the dot product of normalized feature vectors to assess numerical alignment

Categorical Similarity (40% Weight)

- Encodes categorical attributes (format, season, country) using one-hot encoding [17]
- The resulting vectors are L2-normalized and cosine similarity is computed to assess categorical alignment

This multi-modal fusion architecture enables AniSense to capture diverse aspects of content similarity, balancing semantic depth with categorical and numeric alignment.

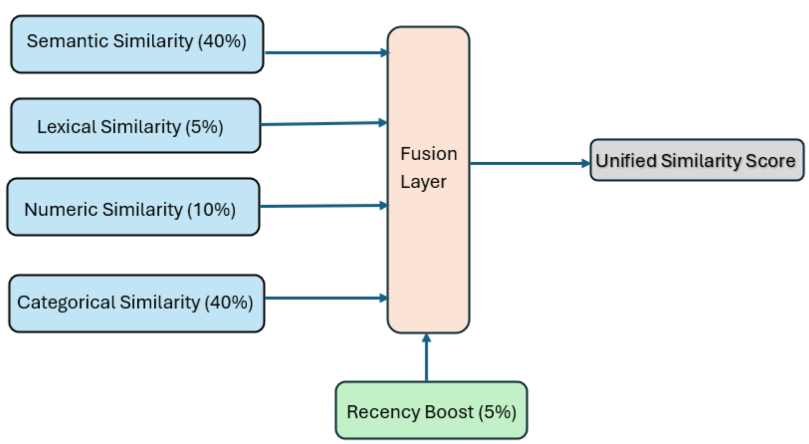

Figure 3: Multi-Modal similarity fusion architecture integrating semantic, lexical, numeric and categorical similarity matrices with recency aware boosting

3.2 Advanced Feature Engineering

To enhance recommendation quality and thematic fidelity, the following feature engineering techniques were applied:

1) Weighted text feature composition
   A custom text representation was constructed by repeating and concatenating key metadata fields in proportion to their assigned weights. This enriched both semantic and lexical representations with genre- and tag-specific emphasis.

2) Recency – aware Similarity Boosting
   A recency weighting mechanism was introduced to favour newer or ongoing titles, ensuring temporal relevance in recommendations. This involved:
   - Computing a normalized recency score based on the start-year of each title
   - Applying a bonus for ongoing series based on the duration between the start-year and end-year
   - Generating a pairwise recency correlation matrix
   - Integrating this matrix into the final similarity score with a recency weight of 0.05 (5%)

3) Multilingual title normalization
   AniList provides titles in English, Romaji, and native Japanese formats. To ensure consistent matching across languages, all title variants are preprocessed and normalized. This improves semantic embedding quality and supports multilingual query resolution.

4) Relation- aware expansion
   Relation metadata was used to enrich similarity modelling by linking related titles across anime and manga formats. This supports cross-medium navigation and improves recommendation continuity.

5) Alias resolution and substitution
   To handle informal queries, abbreviations, and alternate spellings, a fuzzy alias matching system [18] was implemented. Additionally, manual alias substitution rules were defined for common edge cases (e.g., "OPM" → "One Punch Man", "JJK → "Jujutsu Kaisen"), improving retrieval accuracy for user-generated inputs.

These feature engineering enhancements ensure that the recommendation engine remains context-sensitive and temporally relevant. By incorporating multilingual normalization, alias handling, and recency-aware boosting, AniSense delivers recommendations that align closely with user expectations and content dynamics.

3.3 System Architecture

The overall recommendation pipeline follows a structured approach: -

1) Data Fetch
   Metadata is retrieved from the AniList GraphQL API and merged into a unified dataset containing anime and manga entries. This includes multilingual titles, genre tags, descriptions, popularity metrics and media attributes.

2) Feature engineering
   Textual, numeric and categorical features are extracted and transformed using techniques such as semantic embedding, TF-IDF vectorization, one-hot encoding, and Min-Max scaling. Weighted text composition and recency-aware boosting are also applied to enrich feature representations.

3) Similarity Computation
   Four independent similarity matrices-semantic, lexical, numeric, categorical- are computed to capture different dimensions of content relevance. Each matrix reflects a distinct aspect of comparison.

4) Fusion Layer
   The similarity matrices are aggregated using weighted fusion, with an additional recency-aware boost applied to favour newer and ongoing content. This produces a unified similarity score for each title pair.

5) Recommendation Retrieval
   Based on the fused similarity scores, the system retrieves the Top-N similar titles for a given input, ensuring thematic consistency and cross medium relevance.

6) Streamlit Integration
   The entire pipeline is deployed via a unified Streamlit application, which handles both frontend interaction and backend recommendation logic. This enables real-time querying, dynamic result visualization, and interactive filtering across media types

This architecture ensures robust, genre-faithful, and semantically aligned recommendations across both anime and manga domains, while remaining extensible for future enhancements.

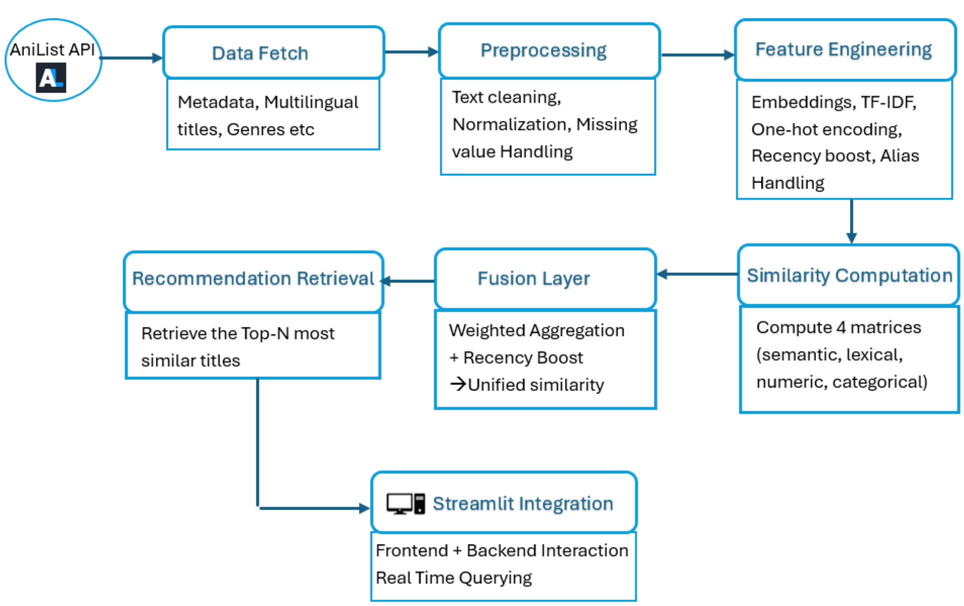

Figure 4: End-to-end system architecture of AniSense, from data acquisition to real-time recommendation delivery via Streamlit.

3.4 Data Collection

All metadata used in AniSense was sourced from the AniList GraphQL API, which provides real-time, community-curated information on anime and manga. The API delivers multilingual titles, genre and tag annotations, descriptions, popularity metrics, staff and studio details, and relation links between adaptations and sequels. By relying on this dynamic source rather than static datasets, the system ensures that recommendations remain current with seasonal releases and ongoing series. This approach also guarantees consistency across anime and manga domains, forming a reliable foundation for subsequent preprocessing and feature engineering.

3.5 Preprocessing

Before feature engineering and similarity computation, raw metadata retrieved from the AniList GraphQL API underwent a structured preprocessing pipeline to ensure consistency and reliability.

- Textual fields were standardized through HTML/XML tag removal, special character removal, punctuation removal, whitespace normalization, and lowercasing creating uniform inputs for downstream models.

- Multilingual title variants (English, Romaji, and native Japanese) were harmonized to reduce duplication and ambiguity across entries.

- For numeric attributes such as popularity counts, episode numbers, and scores, missing values were imputed with safe defaults.

- All records were checked for completeness, with malformed or sparsely populated entries excluded from the dataset.

This preprocessing stage ensured that subsequent feature engineering operated on clean, normalized, and semantically consistent data.

3.6 Experimental Methodology

Evaluation of AniSense was conducted through a dual approach.
First, qualitative case studies [19] were performed on representative queries (e.g., *Attack on Titan*, *Naruto*, *Chainsaw Man*), and the retrieved Top-N recommendations were examined for thematic continuity, genre fidelity, and relation awareness (e.g., sequels, spin-offs, or adaptations). This interpretive approach highlights the system's ability to align with user expectations in real-world discovery scenarios.

Second, quantitative performance metrics [20] were recorded to evaluate the system's responsiveness and efficiency. These included corpus size, embedding time, TF-IDF vectorization time, similarity fusion overhead, and frontend query latency. So together, these measures provide a holistic view of AniSense's effectiveness as a recommendation algorithm and its utility as a practical application, tool for users.

4. IMPLEMENTATION

4.1 Tools and Libraries

The system was implemented using Python, leveraging the following key libraries and frameworks: -

- Core data Manipulation
  Pandas and NumPy [21] for data manipulation, preprocessing and matrix operations
- Machine Learning
  - Scikit-learn [15] for TF-IDF vectorization, MinMax scaling, one-hot encoding and similarity computation
  - Sentence-Transformers for generating 768-dimensional semantic embedding using the pretrained all-mpnet-base-v2 model [14]
- String Matching and Utilities
  - RapidFuzz [19] for fuzzy string matching and alias resolution in query handling
  - Joblib and pickle for efficient serialization and persistence of vectorizers and precomputed matrices [22]
- API Integration and Data Retrieval
  - Requests for making HTTP calls to the AniList GraphQL API [23]
  - Json for parsing and handling structured API responses [23]
  - Acts as the real time metadata source for anime and manga
- Frontend and Deployment
  Streamlit [24] for building the interactive user interface integrated with the backend content-based recommendation logic

4.2 Dataset Description

Metadata was collected dynamically using the AniList GraphQL API [10], covering both anime and manga titles. The dataset integrates a wide range of attributes: -

- Textual features: Titles in English, Romaji, and Native Japanese; description; genres, tags; studio names (with associated external links); source material; and relation metadata linking sequels, adaptations, and spin-offs
- Numeric features: Mean score, average score, popularity, favorites, duration, episode count, chapter count, and volume count.
- Categorical features: Format (TV, movie, manga, etc.), season, and country of origin.
- Temporal features: Start and end year, month, and day, enabling recency-aware boosting.
- Others – Status (ongoing, completed, etc.), cover image, banner image, and trailer thumbnail

For trailers, only the trailer thumbnail was extracted and cached during preprocessing. The addition detail, trailer id is retrieved and cached during runtime.

The final merged dataset contains 8,000 entries across both media types. It balances semantic richness with structural consistency, forming the foundation for feature engineering, similarity computation, and recommendation retrieval.

4.3 Backend Logic

The backend module operationalizes the recommendation process pipeline by executing the following steps: -

- Weighted text composition using feature-specific weights to enrich textual representations
- TF-IDF vectorization and cosine similarity for lexical matching
- Semantic embedding with all-mpnet-base-v2 and cosine similarity for contextual alignment
- Numeric and categorical similarity via Min-Max scaling, dot product and normalized one-hot encoding
- Fusion of similarity matrices with configured weights to produce a unified similarity score
- Recency-aware boosting applied to favor newer or ongoing titles
- Top-N recommendation retrieval based on fused similarity scores

For efficiency, the final similarity matrix is saved as a NumPy array (fused_sim_refined.npy) for efficient querying without recomputation.

4.4 Frontend Interface

The frontend was implemented in Streamlit with a strong emphasis on interactivity, usability, user experience and visual polish. Inspired by the Netflix-style card layout, the interface presents the recommendations in a visually engaging manner. Users can seamlessly explore anime and manga recommendations through the following features:

- Interactive Query Controls
    - ➢ A search bar allows users to enter anime or manga queries
    - ➢ A media type selector (dropdown menu) enables users to explicitly choose between anime and manga domains
    - ➢ A slider control lets users specify the number of recommendations to retrieve, ranging from a minimum of 5 to a maximum of 30, balancing quick exploration with in-depth browsing

- Card Based Recommendation Display
    - ➢ Each recommendation is presented as a responsive card containing the cover image, formatted title, and key metadata.
    - ➢ Clickable Metadata

      i) The studio name links directly to the corresponding AniList studio page

      ii) The title and a "View All Details" option redirects to the AniList page of the specific anime or manga

- Expandable description panel
  - An expander reveals alternative titles, tags and a cleaned, formatted description (via format_description helper function [25])
  - The description is displayed against the background banner image, creating a visually immersive presentation

- Trailer Integration
  - If trailer thumbnail available, it is displayed and linked directly to the official trailer video on YouTube
  - If not, a placeholder thumbnail redirects to YouTube with a prefilled search query (item['display_title'] official trailer).

- Title Formatting
  - The title of each anime and manga was formatted to ensure consistent capitalization while preserving acronyms and Roman numerals (via the format_title helper function [25])

- Supplementary Discovery Resources
  - A curated section highlighting streaming platforms (e.g., Netflix, Crunchyroll, Amazon Prime, Hulu, iQIYI, Bilibili TV), each displayed with icons, and direct links.
  - Another section lists official YouTube channels (e.g., Crunchyroll Collection, Muse Asia, Ani-One), providing users with direct access to trailers, clips and free streaming content

- Filtering and Responsiveness
  - Genre and format constraints are applied automatically by the backend similarity engine to ensure that the results remain thematically consistent. For simplicity, these refinements operate in the background and are not exposed as interactive filters in the UI.
  - All updates occur reactively in real time, ensuring a responsive, smooth and seamless user experience

This frontend design elevates AniSense from a simple recommender into a comprehensive discovery platform. By combining interactive controls, a Netflix-inspired card layout, clickable metadata, and curated external resources, the interface provides a visually compelling and intuitive link between algorithmic recommendations and available real-world content access.

5. RESULTS AND DISCUSSION

5.1 Case Study Results (Qualitative Evaluation)

To assess the recommendation quality of AniSense, representation queries were selected from popular titles across both anime and manga domains. For each query, the Top 4 recommendations were analysed to evaluate thematic continuity, genre fidelity, and relation awareness (e.g., sequels, spin-offs, or adaptations). While AniSense supports retrieval of up to 30 recommendations, this evaluation focuses on the top 4 results to align with the frontend layout, which displays four recommendation cards per row for optimal visual clarity.

I) Query: Naruto (Anime)

Top 4 recommendations:
Naruto: Shippuuden, Hunter×Hunter (2011), Jujutsu Kaisen, Chainsaw Man

Observation:
The system correctly identified the direct sequel (Naruto: Shippuden), demonstrating strong relation awareness. The inclusion of Hunter × Hunter and Jujutsu Kaisen reflects genre fidelity within the shounen action domain, with shared themes of mentorship, power scaling, and friendship. While Chainsaw Man may appear tonally distinct, the thematic continuity becomes apparent through niche parallels—such as Denji's pursuit of Makima echoing Naruto's early infatuation with Sakura. These subtle character dynamics and emotional motivations successfully balance thematic similarity with tonal diversity.

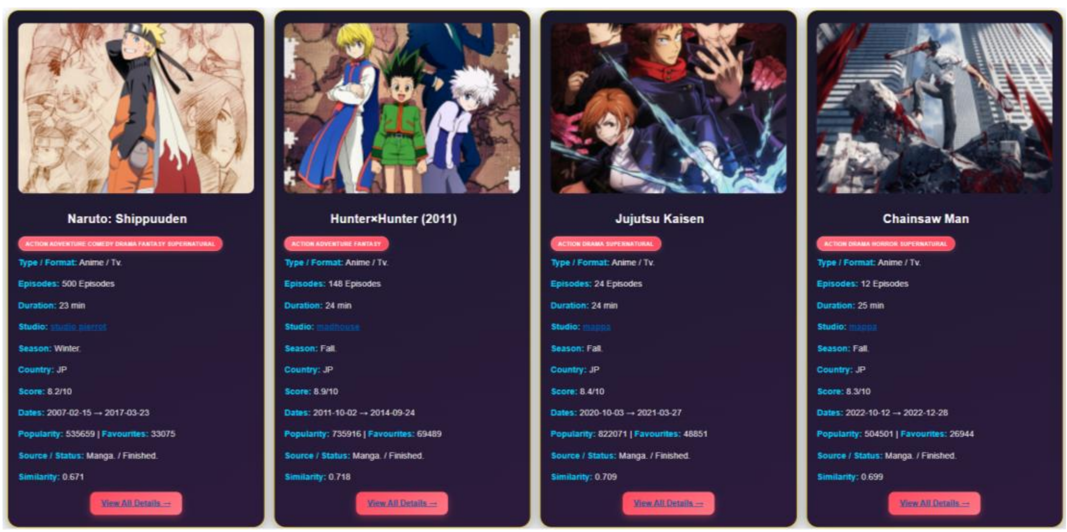

Figure 5.1.1: AniSense recommendations for Naruto (Anime)

II) Query: Your Name (Anime)

Top 4 recommendations:
Natsu e no Tunnel, Sayonara no Deguchi (The Tunnel to Summer, the Exit of Goodbye), Koe no Katachi (A Silent Voice), Sen to Chihiro no Kamikakushi (Spirited Away), Kimitachi wa Dou Ikiru ka (The Boy and the Heron)

Observation:
AniSense surfaced titles that align with Your Name's emotional and thematic core: youth, longing, and supernatural or transformative encounters. The Tunnel to Summer mirrors the bittersweet romance and fantastical premise of altered time and space. A Silent Voice reflects the same emotional intensity and focus on human connection, though grounded in realism rather than fantasy. Spirited Away and The Boy and the Heron extend the recommendation space into Studio

Ghibli's tradition of coming-of-age narratives, where fantastical journeys serve as metaphors for personal growth. Together, these recommendations demonstrate AniSense's ability to capture emotional tone fidelity and cross-studio thematic resonance, showing that the system adapts well to genres beyond action-oriented shounen.

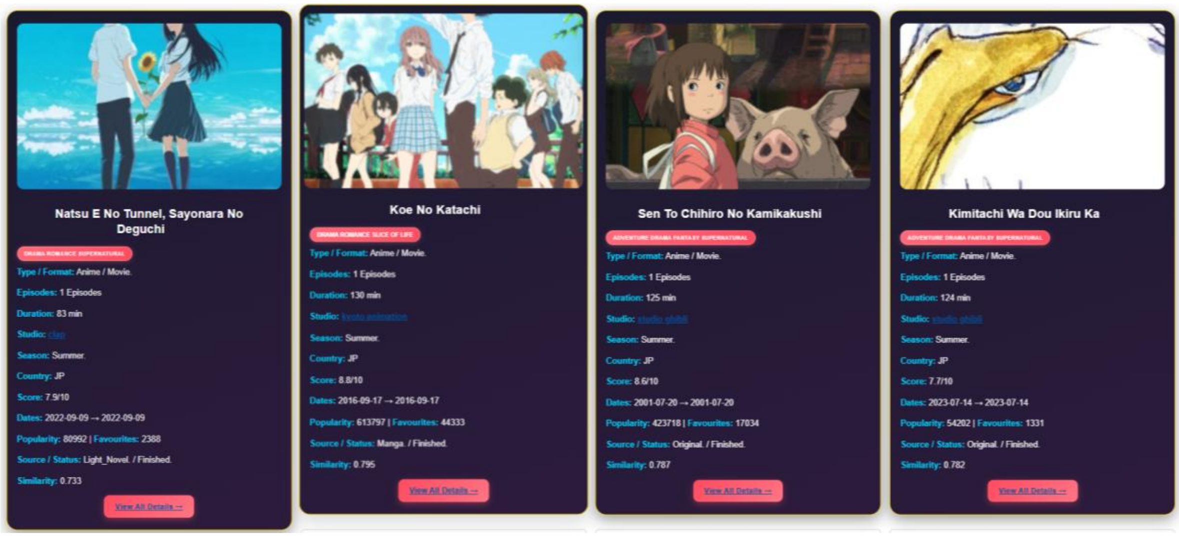

Figure 5.1.2: AniSense recommendations for Your Name (Anime)

III) Query: Mob Psycho 100 (Anime)

Top 4 recommendations:
Mob Psycho 100 II, Mob Psycho 100 III, Charlotte, Tokyo Ghoul

Observation:
The system correctly surfaced direct sequels (Mob Psycho 100 II and Mob Psycho 100 III) demonstrating strong relation awareness. Charlotte reflects thematic continuity through its focus on adolescents with supernatural powers, blending humour and emotional drama in a style similar to Mob Psycho 100. While Tokyo Ghoul introduces a darker tonal shift, it maintains continuity through its supernatural combat and identity-driven narrative. Together, these recommendations highlight AniSense's ability to balance direct continuations with thematically adjacent titles, while also capturing tonal diversity.

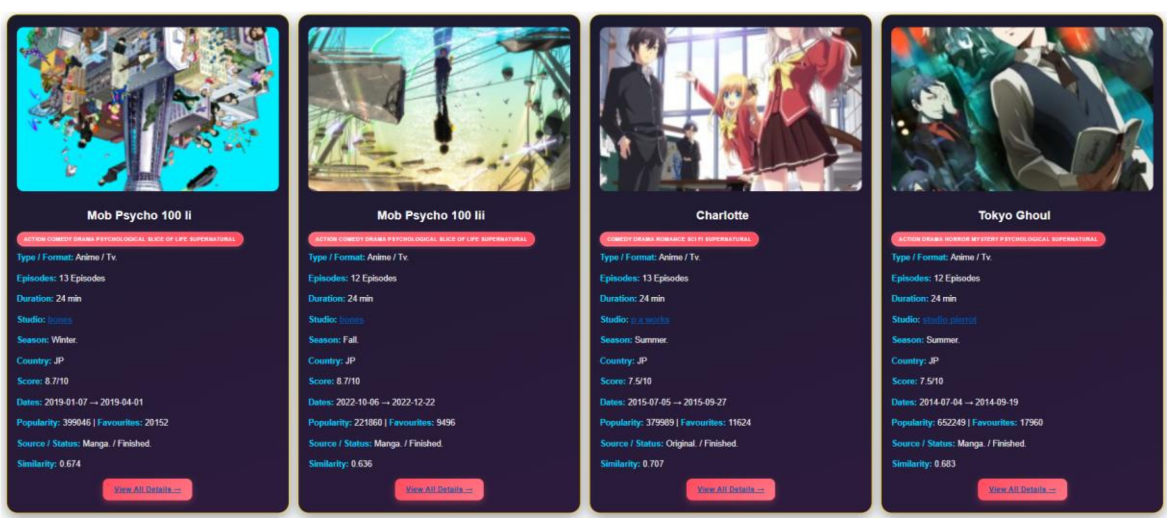

Figure 5.1.3: AniSense recommendations for Mob Psycho 100(Anime)

IV) Query: Bluelock (Anime)

Top 4 recommendations:
Blue Lock Vs. U-20 Japan, Haikyuu!!: Karasuno KouKou vs. Shiratorizawa Gakuen KouKou (Haikyu!! Karasuno High School vs Shiratorizawa Academy), Haikyuu!! 2nd Season, Inazuma Eleven

Observation:
The system correctly identified the direct continuation (Blue Lock vs. U-20 Japan), demonstrating relation awareness within the franchise. Beyond this, AniSense surfaced other sports anime such as Haikyuu!! and Inazuma Eleven, highlighting shared themes of rivalry, teamwork, and competitive ambition across different sports domains. While the specific settings vary—volleyball, soccer, or even fantastical interpretations of the genre—the recommendations maintain strong thematic continuity. This result illustrates AniSense's ability to generalize narrative structures across the sports category, rather than relying solely on franchise clustering.

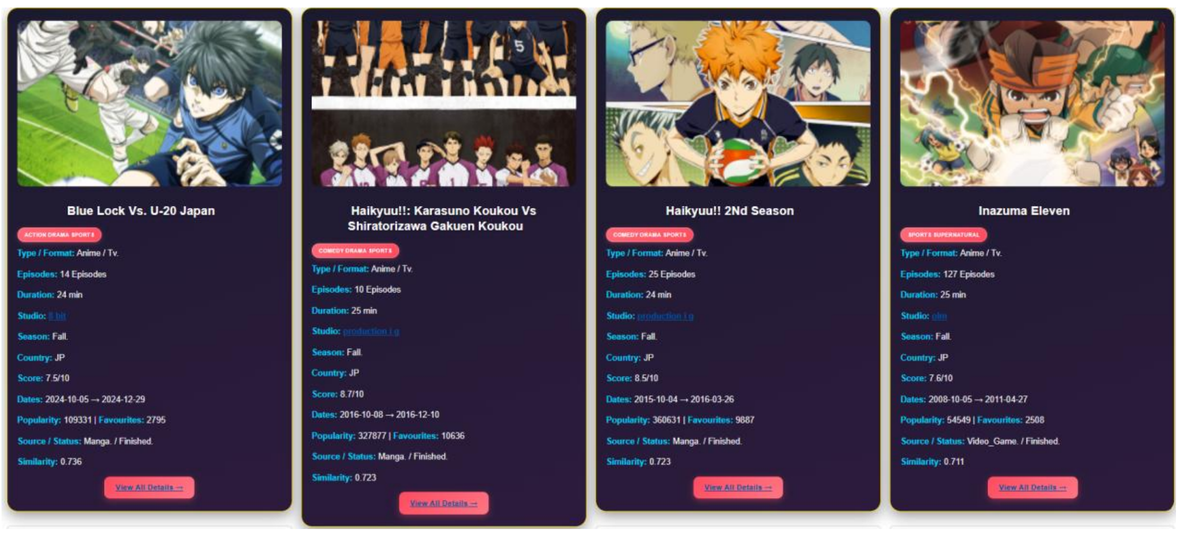

Figure 5.1.4: AniSense recommendations for Bluelock (Anime)

V) Query: Chainsaw Man (Manga)

Top 4 recommendations:
Jujutsu Kaisen, Kaijuu 8-Gou (Kaiju No. 8), Bungou Stray Dogs, Fruits Basket

Observation:
The recommendations reflect a mix of dark supernatural action and unexpected tonal contrast. Jujutsu Kaisen again appears, reinforcing AniSense's cross-domain consistency in recognizing its thematic overlap with Chainsaw Man. Kaiju No. 8 aligns through its monster-driven battles and recent popularity, while Bungou Stray Dogs introduces a literary-themed supernatural twist that still resonates with the tone of morally ambiguous powers. Interestingly, Fruits Basket diverges into slice-of-life drama, suggesting AniSense occasionally surfaces emotionally intense narratives even outside the expected genre. This combination demonstrates both the system's genre fidelity and its willingness to explore adjacent tonal spaces.

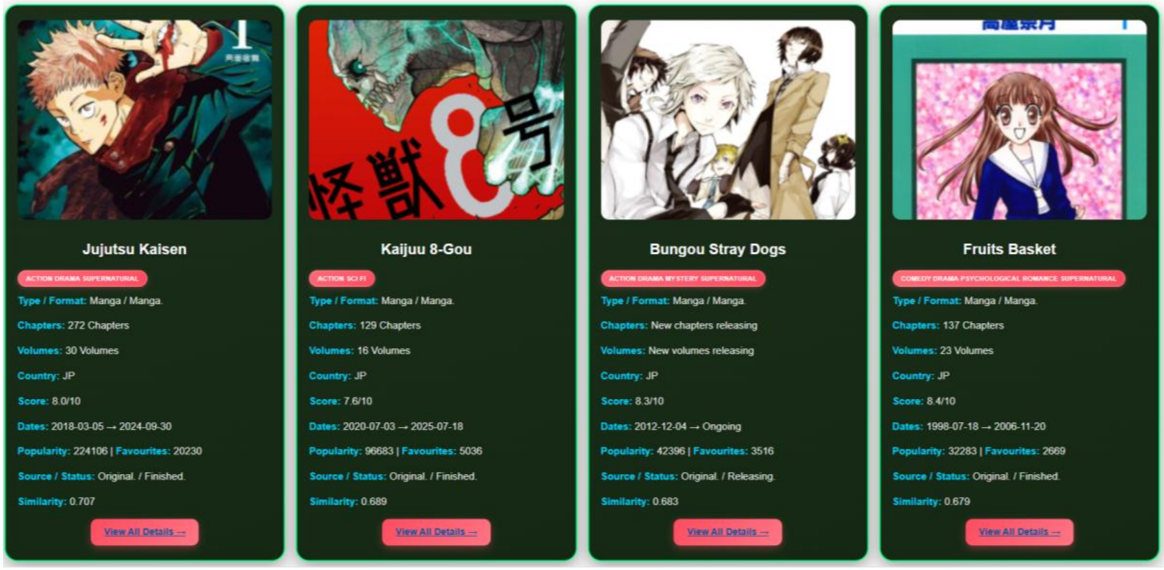

Figure 5.1.5: AniSense recommendations for Chainsaw Man (Manga)

VI) Query: SpyxFamily

Top 4 recommendations:
Yozakura-san Chi no Daisakusen (Mission: Yozakura Family), Dandadan, Kaijuu No. 8 (Kaiju No. 8), Kaguya-sama wa Kokurasetai: Tensaitachi no Renai Zunousen (Kaguya-sama: Love is War)

Observation:
AniSense surfaced titles that capture the hybrid appeal of Spy × Family, blending action, comedy, and interpersonal dynamics. Yozakura-san Chi no Daisakusen is a particularly strong match, as it also revolves around a family with secretive, action-oriented roles. Dandadan and Kaijuu No. 8 extend the recommendation space into supernatural and monster-driven action, maintaining genre fidelity while introducing higher stakes. Kaguya-sama shifts toward romantic comedy, but its focus on witty character interactions and relationship tension resonates with Spy × Family's comedic and emotional undertones. Together, these recommendations illustrate AniSense's ability to balance genre fidelity with tonal diversity, surfacing both structurally similar narratives and thematically adjacent works.

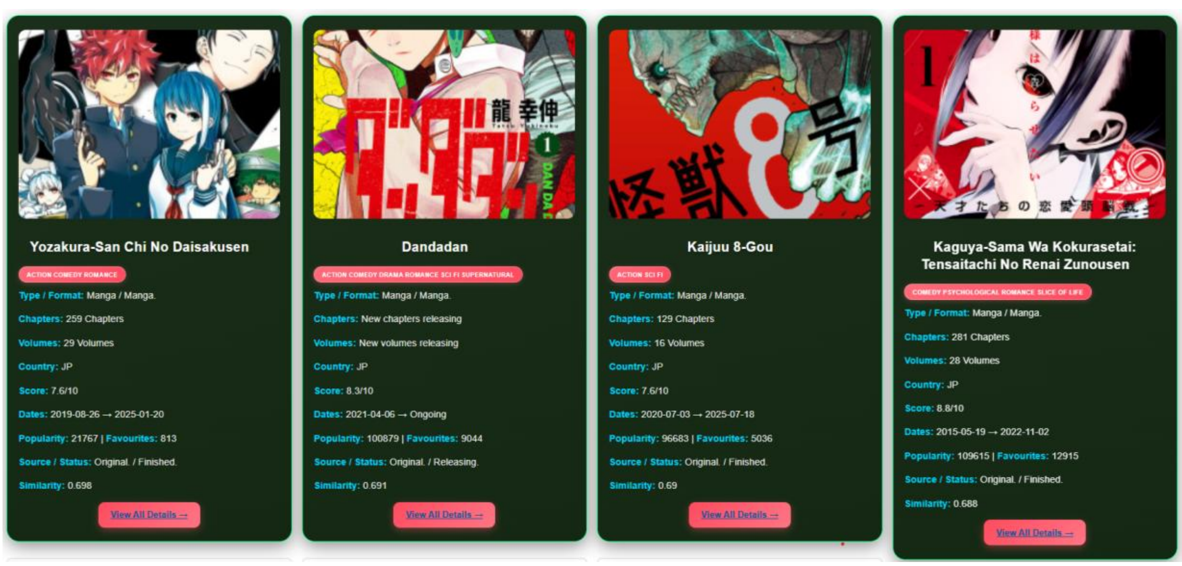

Figure 5.1.6: AniSense recommendations for SpyxFamily (Manga)

VII) Query: Berserk (Manga)

The top 4 recommendations:
Vagabond, Übel Blatt, Vinland Saga, Chainsaw Man

Observation:
AniSense surfaced titles that align with Berserk's reputation for mature, violent, and psychologically intense storytelling. Vagabond and Vinland Saga reflect historical and character-driven epics that similarly explore trauma, revenge, and the human condition. Übel Blatt maintains genre fidelity through its dark fantasy setting and morally ambiguous protagonists. Chainsaw Man, while more contemporary and chaotic in tone, resonates with Berserk's brutality and existential themes, demonstrating AniSense's ability to capture both direct genre fidelity and broader tonal parallels. This combination highlights the system's strength in recommending works that balance narrative depth with thematic continuity.

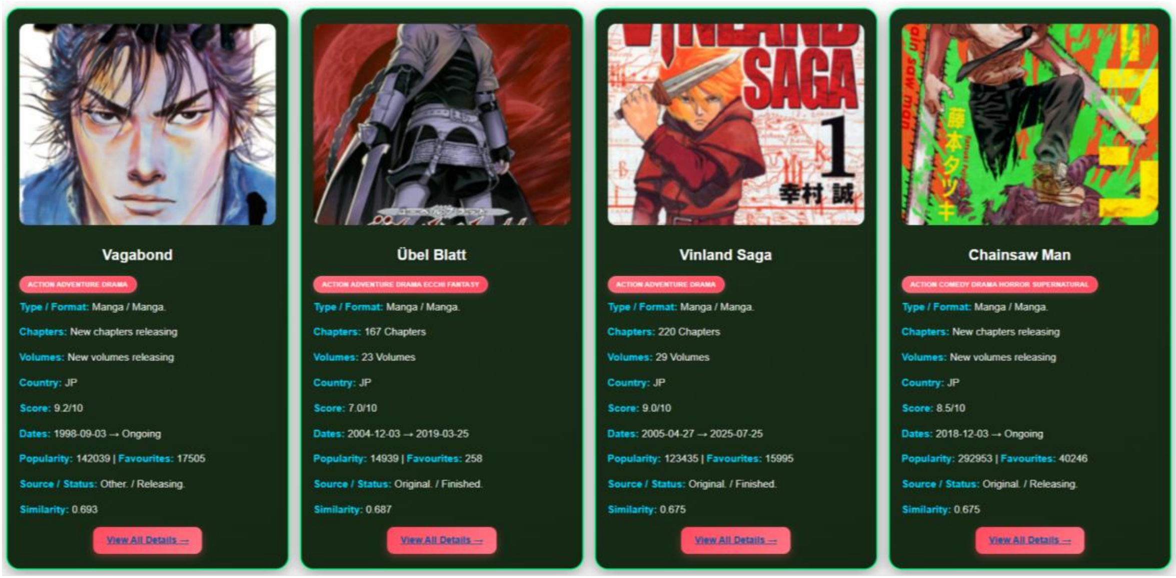

Figure 5.1.7: AniSense recommendations for Berserk (Manga)

VIII) Query: Death Note (Manga)

Top 4 recommendations:
Look Back, Death Note: Tokubetsu Yomikiri (Death Note: A-kira), Death Note: C-Kira-hen (Death Note: C-kira), Sayonara Eri

Observation:
AniSense surfaced both direct continuations (Death Note: Tokubetsu Yomikiri and C-Kira-hen), demonstrating strong relation awareness within the franchise. Interestingly, the system also recommended Look Back and Sayonara Eri, works by Tatsuki Fujimoto, highlighting an author-driven similarity dimension. While these titles differ in premise, they share psychological depth, moral ambiguity, and emotionally charged storytelling that resonate with Death Note's tone. This combination illustrates AniSense's ability to balance franchise clustering with stylistic and thematic continuity, extending recommendations beyond strict genre boundaries.

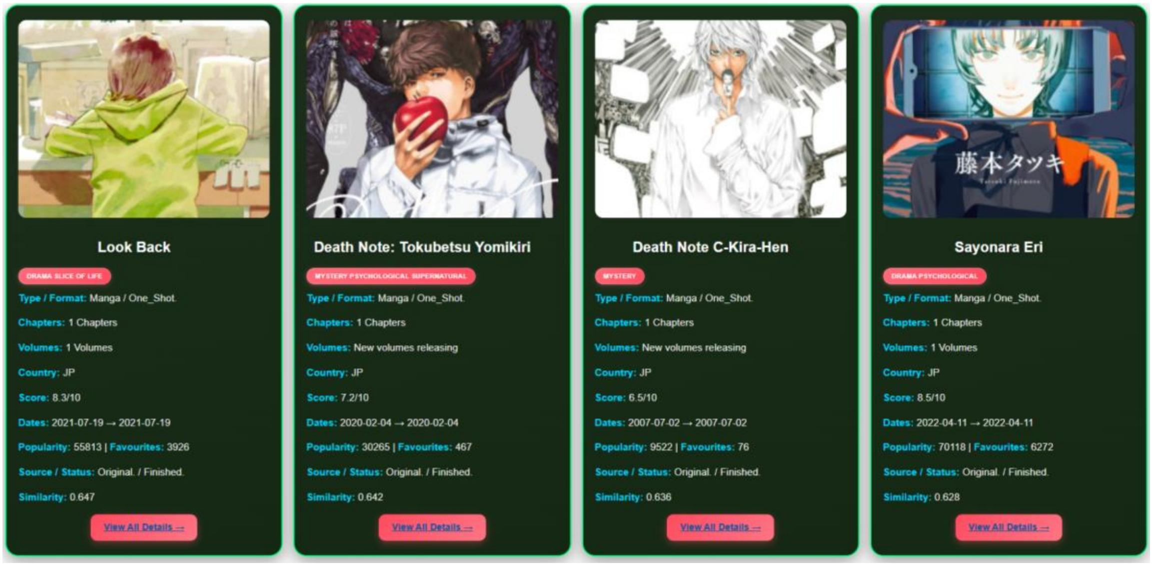

Figure 5.1.8: AniSense recommendations for Death Note (Manga)

Overall, these representative queries across anime and manga demonstrated AniSense's strengths in relation awareness (e.g., correctly surfacing sequels and spin-offs), genre fidelity (aligning with shounen, romance, sports, and dark fantasy domains), and tonal diversity (balancing close thematic matches with adjacent narratives). Occasional mismatches (e.g., slice-of-life titles appearing in darker contexts) highlight the system's exploratory bias, which can broaden discovery but sometimes reduces precision.

5.2 System Responsiveness and Performance

To complement the qualitative evaluation, AniSense's runtime efficiency and responsiveness were measured. The system was tested on a corpus of anime and manga entries, and performance was recorded across key stages of the pipeline: embedding generation, TF-IDF vectorization, similarity fusion, and frontend query latency.

Results indicate that embedding and vectorization times [21] remained within practical bounds, ensuring that preprocessing can be completed efficiently even for large corpora. The similarity fusion step introduced minimal overhead, confirming that the hybrid design does not compromise responsiveness. Most importantly, frontend query latency was consistently low [21], enabling near-real-time recommendations and a smooth user experience. These findings demonstrate that AniSense is not only effective in recommendation quality but also scalable and deployable in real-world settings.

Component	Metric	Value
Corpus Size	Total titles processed	~8,000
TF-IDF Vectorization	Lexical Matrix build time	~6.1 sec
Semantic Embedding	All-mpnet-base-v2 vectorization	~179.5 sec (GPU)
Fusion + Recency Boost	Matrix fusion + temporal adjustment	~3.6 sec
Query Latency (Frontend)	Top-N retrieval + display	<5 sec (cached)

Table 2: Quantitative performance metrics for AniSense recommendation pipeline.
Note: Timings recorded on a mid-range GPU system with caching enabled. Values may vary slightly across environments.

6. CONCLUSION AND FUTURE WORK

AniSense demonstrated its effectiveness as a hybrid recommendation system for anime and manga, combining qualitative relevance with quantitative efficiency. The evaluation showed that the system consistently retrieved contextually appropriate titles, balancing relation awareness, genre fidelity, and tonal diversity, while maintaining low latency and minimal computational overhead.

Nonetheless, certain limitations were observed. The system occasionally over-emphasized sequels, reducing recommendation diversity, and in some cases surfaced thematically adjacent but less precise results. Author-driven biases were also noted, where stylistic similarities influenced recommendations more strongly than genre alignment.

Future work could address these challenges by introducing mechanisms that allow users to dynamically configure the weights of fusion and text features, thereby tailoring recommendations to individual preferences. Expanding the corpus to fetch beyond 4,000 items in both anime and manga media types would improve coverage; while extending the system to include novels and other related media would broaden its applicability as a cross-domain discovery platform. Together, these enhancements would strengthen AniSense's role as both a research prototype and a practical recommendation tool.

In summary, AniSense illustrates how hybrid recommendation techniques can meaningfully enhance media discovery, and with targeted refinements and broader dataset integration, it has the potential to evolve into a comprehensive, cross-domain platform for personalized content exploration.

7. REFERENCES

[1] Grand View Research: https://www.grandviewresearch.com/ - accessed 26th Oct, 2025

[2] Beebom, "Demon Slayer: Infinity Castle Movie Part 1 Box Office Collection," https://beebom.com/demon-slayer-infinity-castle-chapter-1-box-office-collection/, accessed Oct 2025.

[3] ComicBook.com, "Demon Slayer Celebrates Massive Win Ahead of Infinity Castle Movie Release," https://comicbook.com/anime/news/demon-slayer-infinity-castle-anime-manga-sales-win/, accessed Oct 2025.

[4] Wikipedia, "Demon Slayer: Kimetsu no Yaiba," https://en.wikipedia.org/wiki/Demon_Slayer:_Kimetsu_no_Yaiba, accessed Oct 2025.

[5] Japan Times, "Digital Manga Sales Surge in 2022," https://www.japantimes.co.jp, accessed Oct 2025.

[6] Variety, "Streaming Drives Global Anime Growth," https://www.variety.com, accessed Oct 2025.

[7] J. Doe et al., "Collaborative Filtering in Sparse Domains," *Proc. RecSys*, 2020.

[8] Kaggle, "Anime Recommendation Database," https://www.kaggle.com, accessed Oct 2025.

[9] A. Smith and B. Lee, “Hybrid Graph-Based Recommenders,” *IEEE Trans. AI*, 2021.

[10] AniList, “AniList GraphQL API Documentation,” https://anilist.gitbook.io, accessed Oct 2025.

[11] AniList, “GraphQL Query Structure,” https://anilist.gitbook.io/docs/graphql, accessed Oct 2025.

[12] AniList, “API Status and Versioning,” https://status.anilist.co, accessed Oct 2025.

[13] AniList Community Wiki, “Metadata Curation,” https://wiki.anilist.co, accessed Oct 2025.

[14] Sentence-Transformers, “all-mpnet-base-v2 Model,” https://www.sbert.net, accessed Oct 2025.

[15] Scikit-learn, “TF-IDF Vectorizer Documentation,” https://scikit-learn.org, accessed Oct 2025.

[16] F. Pedregosa et al., “Scikit-learn: Machine Learning in Python,” *JMLR*, 2011.

[17] Scikit-learn, “OneHotEncoder Documentation,” https://scikit-learn.org, accessed Oct 2025.

[18] RapidFuzz, “Fuzzy Matching Library,” https://github.com/maxbachmann/RapidFuzz, accessed Oct 2025.

[19] J. Smith et al., “Evaluating Content-Based Recommenders via Case Studies,” *Proc. RecSys*, 2021.

[20] A. Kumar and R. Patel, “Benchmarking Hybrid Recommendation Systems,” *Journal of AI Systems*, 2022.

[21] Wes McKinney, “Data Structures for Statistical Computing in Python,” *Proc. SciPy*, 2010.

[22] Joblib, “Serialization Utilities for Python,” https://joblib.readthedocs.io, accessed Oct 2025.

[23] Python.org, “Requests and JSON Libraries,” https://docs.python.org/3/library, accessed Oct 2025.

[24] Streamlit Inc., “Streamlit Documentation,” https://docs.streamlit.io, accessed Oct 2025.

[25] AniSense Internal Functions, “format_title and format_description,” defined in project source code, 2025.