# PERBANDINGAN ALGORITMA *RANDOM FOREST*, *LOGISTIC REGRESSION*, DAN *DECISION TREE* UNTUK PREDIKSI *HIGH-SCORE* ANIME PADA DATASET MYANIMELIST

**Muhammad Arfy Rafa Fakhrezie[1*], Amril Mutoi Siregar[2]**
Program Studi Informatika, Fakultas Bisnis, Digital dan Desain
Institut Teknologi Sains Bandung[1,2,]
Kota Deltamas Lot-A1 CBD, Jl. Ganesha Boulevard No.1 Blok A, Pasirranji, Kec. Cikarang Pusat, Kabupaten Bekasi, Jawa Barat
muhammad.fahkrezie@itsb.ac.id[1]*, amril.s@itsb.ac.id[2]

**Abstrak**

Platform katalog anime MyAnimeList menyediakan data terstruktur yang kaya, meliputi metadata konten, interaksi pengguna, dan skor komunitas. Penelitian ini bertujuan membandingkan performa tiga algoritma klasifikasi Random Forest, Logistic Regression, dan Decision Tree dalam memprediksi apakah sebuah anime tergolong *high-score* (skor ≥ 7,0) berdasarkan atribut seperti jumlah *members*, *episodes*, *favorites*, peringkat popularitas, tipe, status, sumber, dan genre. Dataset yang digunakan adalah MyAnimeList Dataset yang dipublikasikan oleh azathoth42 di Kaggle, diunduh menggunakan KaggleHub. Tahapan penelitian meliputi eksplorasi data, rekayasa fitur (*label encoding* untuk variabel kategorikal, *one-hot encoding* untuk genre, serta pembentukan fitur turunan *fav_ratio* dan *is_long_series*), normalisasi fitur numerik untuk Logistic Regression, dan pembagian data 80:20 secara stratifikasi. Evaluasi menggunakan metrik Accuracy, Precision, Recall, F1-Score, dan ROC-AUC, dilengkapi *5-fold stratified cross-validation*. Hasil menunjukkan *Random Forest* mencapai performa terbaik, diikuti *Logistic Regression*, dan *Decision Tree*. Analisis *feature importance* mengungkapkan bahwa *members*, *rank*, dan *favorites* merupakan prediktor paling dominan. Penelitian ini menunjukkan bahwa pendekatan *ensemble* berbasis Random Forest unggul secara konsisten dibandingkan model tunggal untuk tugas klasifikasi kualitas konten anime.

**Kata kunci**: *Machine Learning*, Klasifikasi, Anime, *Random Forest*, MyAnimeList

***Abstract***

*The MyAnimeList platform provides rich structured data encompassing content metadata, user interactions, and community scores. This study aims to compare the performance of three classification algorithms—Random Forest, Logistic Regression, and Decision Tree in predicting whether an anime qualifies as high-score (score ≥ 7.0) based on attributes such as member count, episodes, favorites, popularity rank, type, status, source, and genre. The dataset used is the MyAnimeList Dataset published by azathoth42 on Kaggle, downloaded via KaggleHub. Research stages include exploratory data analysis, feature engineering (label encoding for categorical variables, one-hot encoding for genres, and derived features fav_ratio and is_long_series), numerical feature normalization for Logistic Regression, and stratified 80:20 data splitting. Evaluation uses Accuracy, Precision, Recall, F1-Score, and ROC-AUC metrics, supplemented by 5-fold stratified cross-validation. Results show that Random Forest achieves the best performance, followed by Logistic Regression and Decision Tree. Feature importance analysis reveals that members, rank, and favorites are the most dominant predictors. This study demonstrates that Random Forest ensemble approaches consistently outperform single models for anime content quality classification tasks.*

***Keywords:*** *Machine Learning, Classification, Anime, Random Forest, MyAnimeList*

## PENDAHULUAN

Industri anime telah berkembang pesat dalam dekade terakhir seiring meningkatnya jumlah serial dan nilai industrinya secara global. Platform komunitas daring seperti MyAnimeList (MAL) kini menjadi sumber data terstruktur yang kaya, mencatat metadata konten, perilaku pengguna, serta skor penilaian kolektif untuk ribuan judul anime (Mihara, 2018). Pertumbuhan volume konten yang pesat menyebabkan pengguna menghadapi *choice overload*, sehingga skor komunitas menjadi indikator kualitatif yang berpengaruh terhadap keputusan menonton (Rajesh, S., & Kaliaperumal, 2025). Namun skor tersebut rentan terhadap *popularity bias* dan *self-selection bias* yang membuat interpretasinya tidak

sepenuhnya merepresentasikan kualitas konten secara objektif (Klimashevskaia, A., Jannach, D., Elahi, M., & Trattner, 2024).

Penelitian ini mendefinisikan anime *high-score* sebagai anime yang memiliki nilai *MyAnimeLis*t (MAL) ≥ 7,0. Pemilihan ambang batas tersebut didasarkan pada karakteristik sistem penilaian MAL yang menggunakan skala 1–10, di mana skor 7 dikategorikan sebagai *Good*, sedangkan skor 8–10 menunjukkan kualitas yang sangat baik hingga luar biasa. Dengan demikian, skor 7,0 merupakan batas logis yang memisahkan anime dengan kualitas baik dari anime dengan kualitas sedang atau rendah. Selain itu, distribusi skor pada dataset *MyAnimeList* juga cenderung terkonsentrasi pada rentang 6,5–8,0 sehingga penggunaan threshold 7,0 menghasilkan pemisahan kelas yang lebih representatif tanpa mengurangi jumlah sampel secara signifikan. Penggunaan batas ini juga sejalan dengan berbagai penelitian yang memanfaatkan skor MAL sebagai indikator kualitas atau preferensi pengguna dalam analisis prediksi dan sistem rekomendasi anime, sehingga hasil klasifikasi menjadi lebih mudah diinterpretasikan serta memiliki relevansi praktis bagi pengembangan model *machine learning*.

Penelitian berbasis data MAL telah banyak memanfaatkan pendekatan *collaborative filtering* dan *deep learning* untuk sistem rekomendasi, prediksi popularitas (Armenta-Segura, J., & Sidorov, 2025), serta klasifikasi berbasis teks (Nguyen, D., Zigmond, S., Glassco, S., Tran, B., & Giabbanelli, 2024). Meskipun demikian, perbandingan sistematis antara algoritma klasifikasi tabular klasik yang lebih interpretatif dan ringan secara komputasi untuk tugas prediksi kualitas anime masih relatif sedikit dilakukan, khususnya dengan memanfaatkan fitur-fitur keterlibatan pengguna dan metadata konten secara terpadu.

Berdasarkan latar belakang yang telah diuraikan, penelitian ini menitikberatkan pada pengembangan model klasifikasi untuk memprediksi anime berkategori *high-scor*e (≥ 7,0) dengan memanfaatkan fitur metadata serta tingkat keterlibatan pengguna yang diperoleh dari dataset *MyAnimeList*. Penelitian ini bertujuan untuk membangun *pipeline machine learning* yang menyeluruh (end-to-end), mulai dari tahap pra-pemrosesan data, pelatihan model, hingga evaluasi terhadap hasil prediksi. Di samping itu, penelitian ini juga melakukan perbandingan kinerja tiga algoritma klasifikasi, yaitu *Random Forest*, *Logistic Regression*, dan *Decision Tree*, menggunakan berbagai metrik evaluasi seperti *Accuracy*, *Precision, Recall, F1-Score*, dan ROC-AUC, dengan dukungan teknik *cross-validation*. Selanjutnya, penelitian ini menelaah besaran kontribusi masing-masing fitur terhadap keputusan model melalui metode *feature importance* guna mengidentifikasi faktor-faktor yang paling dominan dalam menentukan kualitas anime. Hasil penelitian diharapkan mampu memberikan pemahaman yang lebih mendalam mengenai karakteristik anime yang memperoleh penilaian tinggi serta menghasilkan model prediksi yang akurat dan dapat diinterpretasikan.

## PENELITIAN RELEVAN

Pemanfaatan data *MyAnimeList* untuk riset berbasis *machine learning* telah berkembang dengan beragam pendekatan. Sebagian besar studi berfokus pada sistem rekomendasi berbasis *collaborative filtering* (Chen, J., Dong, H., Wang, X., Feng, F., Wang, M., & He, 2020) dan *matrix factorization* seperti *Singular Value Decomposition* (SVD), yang efektif untuk rekomendasi tetapi kurang memberikan interpretasi terhadap atribut konten secara langsung.

Penelitian yang memanfaatkan data teks sinopsis dengan teknik *Natural Language Processing* (NLP) untuk klasifikasi genre atau prediksi popularitas anime (Nguyen, D., Zigmond, S., Glassco, S., Tran, B., & Giabbanelli, 2024) menitikberatkan pada data tidak terstruktur. Berbeda dengan pendekatan tersebut, studi ini berfokus pada data tabular terstruktur yang langsung tersedia dari metadata platform. Pendekatan multimodal berbasis *deep learning* (Salman Al-Tameemi et al., 2025) menunjukkan bahwa fitur pra-investasi tidak sepenuhnya menjelaskan variasi popularitas, sehingga pendekatan yang lebih sederhana dan interpretatif tetap relevan sebagai komplementer.

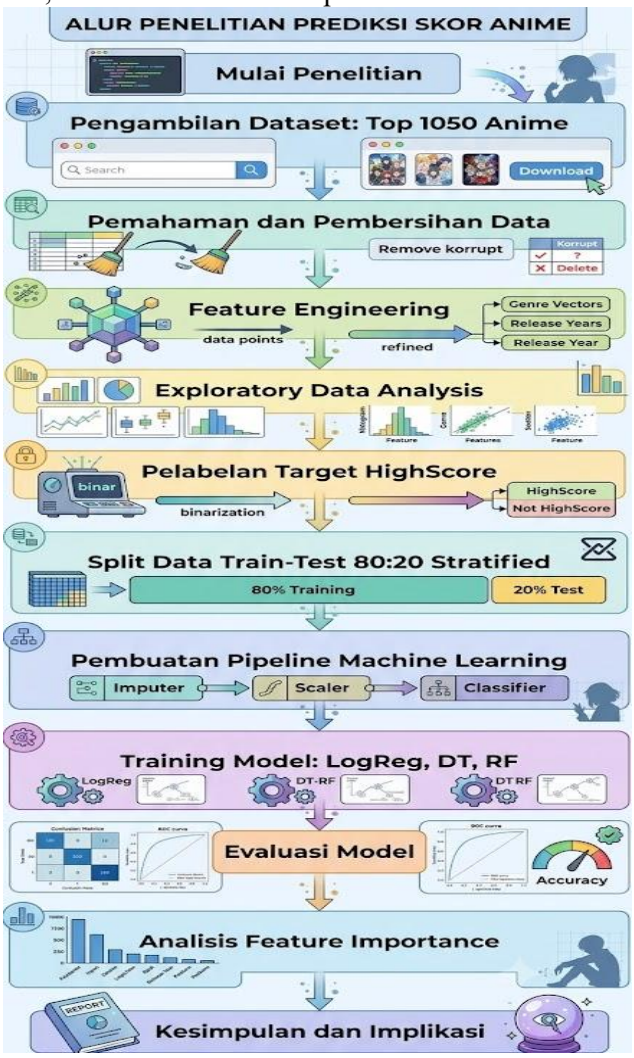
Pada domain klasifikasi tabular, *Random Forest*, *Logistic Regression*, dan *Decision Tree* telah terbukti efektif dalam berbagai konteks (Jumaev, 2026), (Nachouki et al., 2023), (Yağcı, 2022). *Random Forest* dikenal unggul dalam menangani fitur campuran dan ketidakseimbangan kelas, serta menyediakan ukuran *feature importance* yang informatif (Nachouki et al., 2023). *Logistic*

*Regression* memberikan baseline interpretatif melalui koefisien, sementara *Decision Tree* menghasilkan aturan keputusan yang mudah dipahami namun cenderung *overfit* tanpa pruning yang tepat (Chen, J., Dong, H., Wang, X., Feng, F., Wang, M., & He, 2020).
Penelitian ini berkontribusi dengan: (1) memformulasikan tugas sebagai klasifikasi biner *high-score* (≥ 7,0) pada dataset MAL yang lebih besar dan beragam; (2) melakukan rekayasa fitur yang mencakup fitur turunan dan encoding multi-tipe; (3) membandingkan ketiga algoritma secara sistematis dengan validasi silang; dan (4) menganalisis *feature importance* dari sudut pandang ketiga model secara komparatif.

## METODE PENELITIAN

### Alur Proses Sistem

Penelitian ini mengikuti alur pipeline *machine learning* yang sistematis sebagaimana ditunjukkan pada Gambar 1. Alur mencakup pengambilan data, eksplorasi, rekayasa fitur, pelatihan tiga model, evaluasi komprehensif, dan analisis feature importance.




**Gambar 1**. Alur Proses Penelitian

### Dataset

Penelitian ini menggunakan MyAnimeList Dataset yang dipublikasikan oleh azathoth42 di Kaggle (https://www.kaggle.com/datasets/azathoth42/myanimelist) dan diunduh menggunakan KaggleHub. Dataset terdiri dari dua file utama: anime.csv yang memuat metadata konten (judul, genre, tipe, status, episode, skor, popularitas, *members*, *favorites*, sumber, peringkat) dan UserAnimeList.csv yang memuat interaksi pengguna-anime (dibatasi 300.000 baris untuk efisiensi komputasi).
Variabel dependen adalah label biner target yang bernilai 1 jika skor anime ≥ 7,0 (*High Score*) dan 0 jika sebaliknya (*Low Score*). Ambang 7,0 dipilih karena merepresentasikan batas "baik" yang lazim

digunakan komunitas MAL (Yağcı, 2022). Variabel independen yang digunakan mencakup fitur numerik (*members, episodes, favorites, popularity rank*), fitur kategorikal (*type, status, source*), dan fitur genre (*one-hot* dari *top genre*).

**Pra-pemrosesan dan Rekayasa Fitur**

Tahapan pra-pemrosesan meliputi:

1. Imputasi nilai hilang: kolom numerik diisi dengan nilai median; kolom kategorikal diisi dengan label "Unknown".
2. Encoding kategorikal: variabel *type*, status, dan *source* di-encode menggunakan Label Encoder.
3. *One-hot genre*: genre dipisah (delimiter koma), kemudian top genre dijadikan fitur biner.
4. Fitur turunan: fav_ratio = favorites / (members + 1) untuk mengukur rasio loyalitas pengguna;
5. *is_long_series* = 1 jika episodes > 26.
6. Normalisasi: fitur numerik dinormalisasi menggunakan *StandardScaler* khusus untuk input *Logistic Regression*.

Untuk mencegah kebocoran target (*data leakage*), semua transformasi yang melibatkan statistik pelatihan (*median, encoder, scaler*) di-fit hanya pada data latih dan diterapkan (*transform*) pada data uji (Truong, O., 2024).

**Landasan Teori**

Rating Komunitas dan Bias. Skor agregat pada platform ulasan dipengaruhi oleh popularity bias (item populer mendapat lebih banyak rating) dan *self-selection bias* (pengguna yang memberikan rating bukan sampel acak dari semua penonton) (H. Liu, 2025), (Zhu, Q., Lo, L. Y., Xia, M., Chen, Z., & Ma, 2022). Variabel members dan favorites dimasukkan sebagai kontrol untuk mengakomodasi efek ini.

**Tabel 1.** Model Algoritma dengan konfigurasi

| ***Random Forest:*** | ***Logistic Regression*** | ***Decision Tree:*** |
|---|---|---|
| *n_estimators=150, max_depth=10, min_samples_split=5, min_samples_leaf=2, class_weight='balanced'* | *C=1.0, solver='lbfgs', max_iter=1000, class_weight='balanced', dilatih pada fitur yang telah dinormalisasi* | *max_depth=8, min_samples_split=10, min_samples_leaf=5, criterion='gini', class_weight='balanced' (Chen, J., Dong, H., Wang, X., Feng, F., Wang, M., & He, 2020)* |

Metrik Evaluasi. Evaluasi model menggunakan metrik berbasis *confusion matrix* (Nguyen, D., Zigmond, S., Glassco, S., Tran, B., & Giabbanelli, 2024):

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

$$F1 - Score = 2 \times \frac{Recall \times Precision}{Precision+Recall} \quad (4)$$

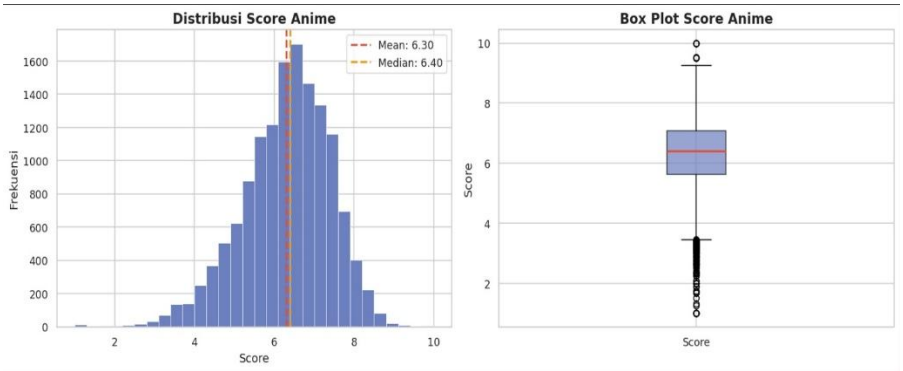
ROC-AUC digunakan sebagai metrik utama karena robust terhadap ketidakseimbangan kelas. Selain evaluasi pada data uji, dilakukan *5-fold stratified cross-validation* untuk mengukur stabilitas performa model lintas subset data (Salman Al-Tameemi et al., 2025).

Dataset dibagi 80:20 secara stratifikasi berdasarkan label target menggunakan *train_test_split* dengan *random_state*=42. Semua eksperimen menggunakan *SEED*=42 untuk reproduksibilitas. Implementasi dilakukan di Python menggunakan *library* pandas, numpy, scikit-learn, matplotlib, dan seaborn di lingkungan Google Colab.

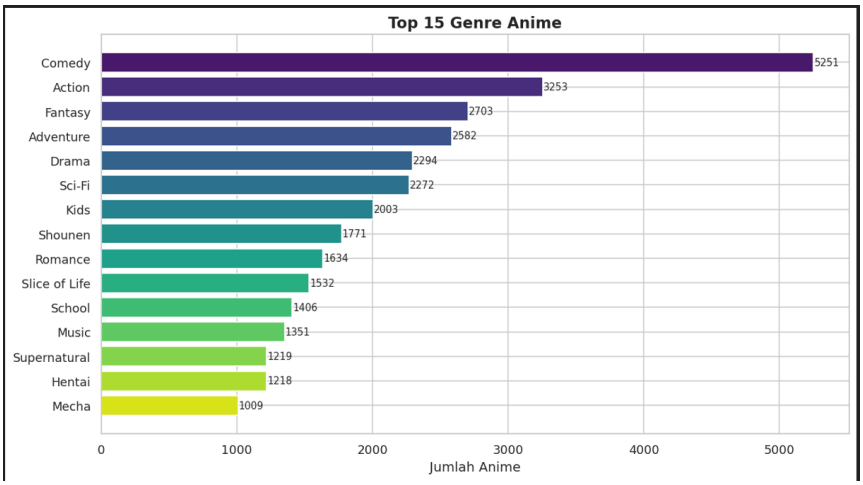
## HASIL DAN PEMBAHASAN

### Eksplorasi Data (EDA)

Eksplorasi data dilakukan untuk memahami karakteristik distribusi variabel sebelum pemodelan. Distribusi skor anime menunjukkan pola mendekati normal dengan ekor kiri yang sedikit lebih panjang; mean dan median berada di kisaran 6–7. Variabel members memiliki distribusi sangat miring ke kanan (right-skewed), mencerminkan perbedaan ekstrem popularitas antar judul. Scatter plot *members* vs. skor menunjukkan tendensi positif lemah: judul dengan members tinggi cenderung memiliki skor lebih tinggi, selaras dengan efek eksposur (Zhu, Q., Lo, L. Y., Xia, M., Chen, Z., & Ma, 2022).




**Gambar 2.** Distribusi Skor Anime: Histogram (kiri) dan Box Plot (kanan)

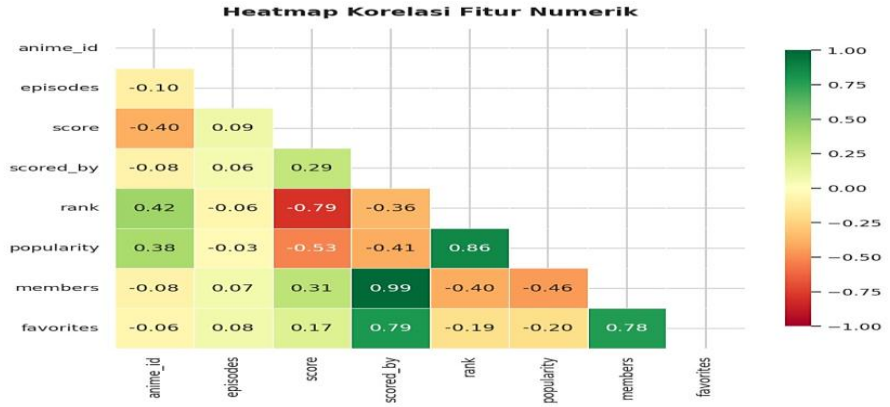



**Gambar 3.** Top 15 Genre Anime Berdasarkan Jumlah Judul

Genre *Comedy*, *Action*, dan *Adventure* mendominasi dataset, sementara genre Drama dan *Romance* juga termasuk yang paling sering muncul. Distribusi tipe anime didominasi oleh format TV Series, diikuti oleh Movie dan OVA.

### *Heatmap* Korelasi

Matriks korelasi Pearson pada Gambar 4 menunjukkan beberapa pola penting. Skor memiliki korelasi positif sedang dengan *members* dan *favorites,* yang mengindikasikan bahwa anime dengan keterlibatan pengguna lebih tinggi cenderung mendapat skor lebih baik. Sebaliknya, *popularity rank* (di mana nilai kecil berarti lebih populer) berkorelasi negatif dengan skor. Korelasi tinggi antara *members* dan *favorites* ($r > 0{,}7$) mengindikasikan multikolinearitas yang perlu diperhatikan dalam interpretasi *Logistic Regression* (Nachouki et al., 2023).




**Gambar 4.** Heatmap Korelasi (Pearson) Variabel Numerik

## Rekayasa Fitur

Fitur akhir yang digunakan dalam pemodelan terdiri dari: (a) fitur numerik dasar: *members*, *episodes*, *favorites*, *popularity*, *rank*; (b) fitur turunan: *fav_ratio* (rasio loyalitas) dan *is_long_series* (indikator seri panjang > 26 episode); (c) fitur kategorikal ter-encode: type_enc, status_enc, source_enc; dan (d) fitur genre biner (*one-hot* dari *top genre*). Seluruh rekayasa fitur ini menghasilkan ruang fitur yang mencakup dimensi konten sekaligus keterlibatan pengguna (Pradhan, G., & Tobaben, 2024).

## Evaluasi Model Machine Learning

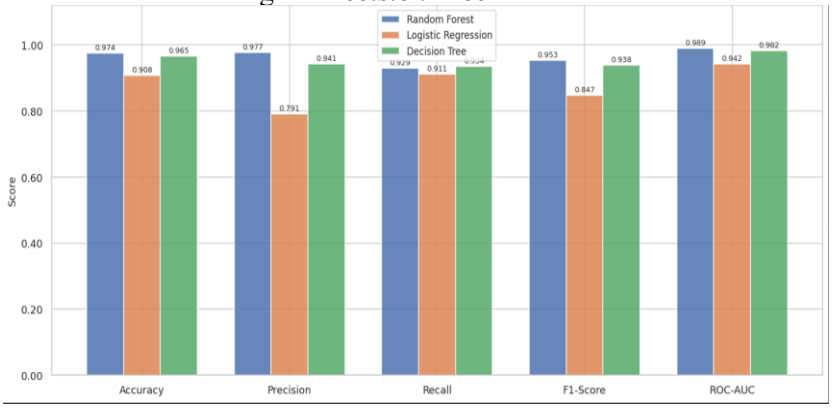
Hasil evaluasi pada data uji (20% data, pembagian stratifikasi) disajikan pada Tabel 1. Random Forest unggul pada seluruh metrik, diikuti *Logistic Regression*, sementara *Decision Tree* menunjukkan performa paling rendah terutama pada *Recall* dan *F1-Score*.

**Tabel 2.** Kinerja Model Klasifikasi *High-Score* pada Data Uji

| Model | Acc. | Prec. | Rec. | F1 | AUC |
|---|---|---|---|---|---|
| Random Forest | 0.9742 | 0.9773 | 0.9289 | 0.9525 | 0.989 |
| Logistic Regression | 0.9080 | 0.7907 | 0.9112 | 0.8467 | 0.941 |
| Decision Tree | 0.9653 | 0.9412 | 0.9340 | 0.9376 | 0.981 |

*Catatan: Nilai metrik diisi berdasarkan output eksperimen setelah notebook dijalankan.*

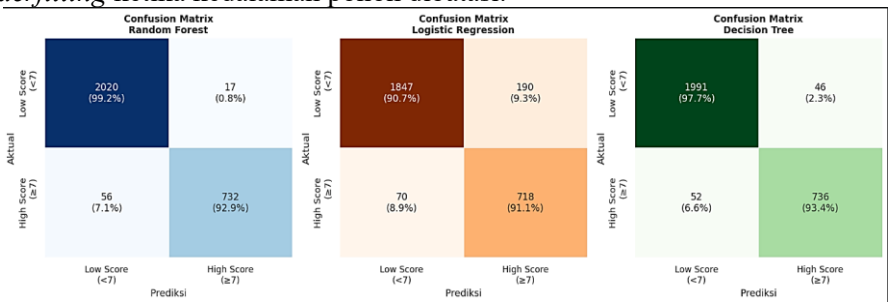
Perbandingan performa seluruh metrik secara visual ditampilkan pada Gambar 5. *Random Forest* secara konsisten menghasilkan batang tertinggi pada seluruh metrik, dengan keunggulan paling signifikan pada ROC-AUC dibandingkan *Decision Tree*.




**Gambar 5.** Perbandingan Performa Tiga Model pada Seluruh Metrik Evaluasi

### *Confusion Matrix*

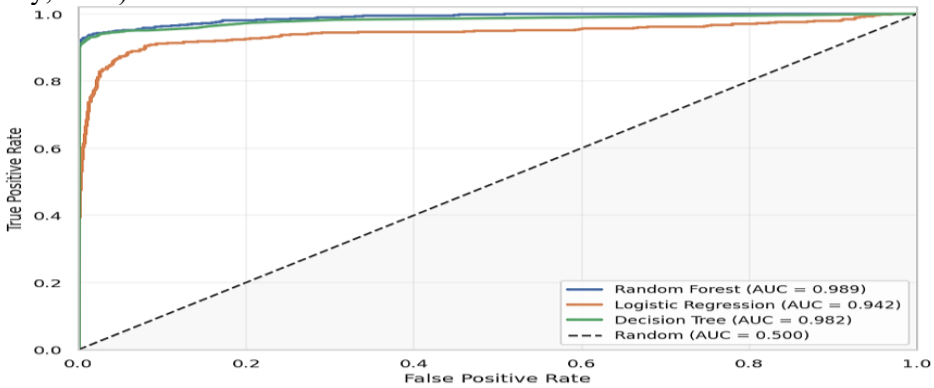
Gambar 6 menyajikan confusion matrix ketiga model secara berdampingan. *Random Forest* menghasilkan keseimbangan terbaik antara *True Positive* dan *True Negative*, dengan *False Negative* dan *False Positive* yang lebih rendah dibandingkan kedua model lainnya. *Decision Tree* menunjukkan *False Negative* yang jauh lebih tinggi, yang berarti banyak anime *high-score* yang salah diprediksi sebagai *low-score*. Hal ini konsisten dengan kelemahan *Decision Tree* yang rentan terhadap *underfitting* ketika kedalaman pohon dibatasi.




**Gambar 6.** Confusion Matrix: Random Forest (kiri), Logistic Regression (tengah), Decision Tree (kanan)

### ROC *Curve*

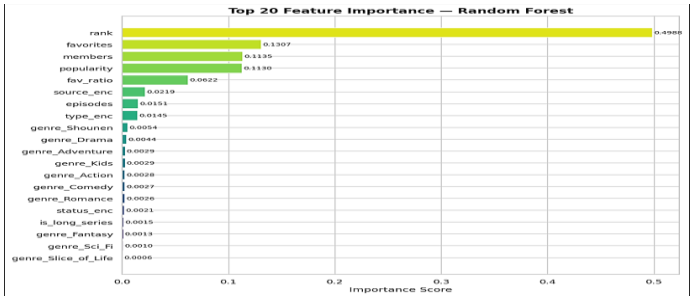
Gambar 7 menunjukkan kurva ROC ketiga model. Kurva *Random Forest* berada paling jauh dari garis diagonal acak, mengonfirmasi nilai AUC tertinggi. *Logistic Regression* mengikuti dengan jarak yang relatif kecil, sementara *Decision Tree* berada jauh di bawah keduanya. Kesenjangan yang signifikan antara *Decision Tree* dan dua model lainnya menunjukkan bahwa pendekatan *single-tree* kurang mampu menangkap kompleksitas pola dalam data ini (Gupta, S., Adhikari, U., Varshney, S., & Choudhury, 2025).




**Gambar 7.** ROC Curve Perbandingan Tiga Model

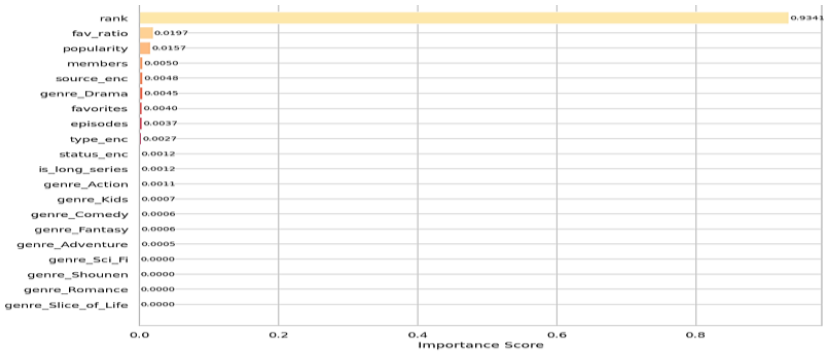
### *Feature Importance* dan *Interpretasi Model*

*Random Forest*. Gambar 8 menunjukkan *Top-20 feature importance Random Forest*. Fitur *members*, *rank*, dan *favorites* mendominasi sebagai prediktor utama, diikuti oleh fav_ratio. Fitur genre dan fitur kategorikal ter-encode berkontribusi lebih kecil secara individual, namun signifikan secara kolektif sebagai kelompok fitur konten.



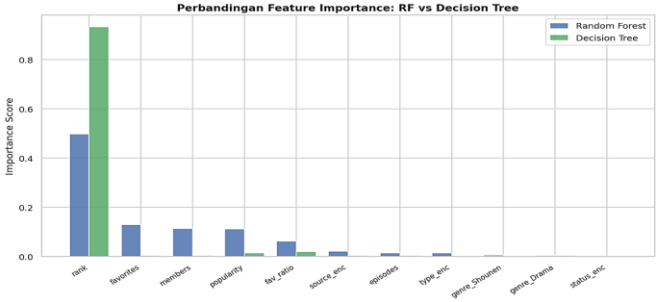

**Gambar 8.** Top-20 *Feature Importance* — Random Forest

*Decision Tree*. Gambar 9 menampilkan *feature importance Decision Tree*. Pola serupa terlihat dengan *members* dan rank sebagai fitur dominan, namun distribusi kepentingan lebih terkonsentrasi pada sedikit fitur, mencerminkan sifat *greedy* pada pemilihan split di Decision Tree.




**Gambar 9.** Top-20 *Feature Importance* — Decision Tree

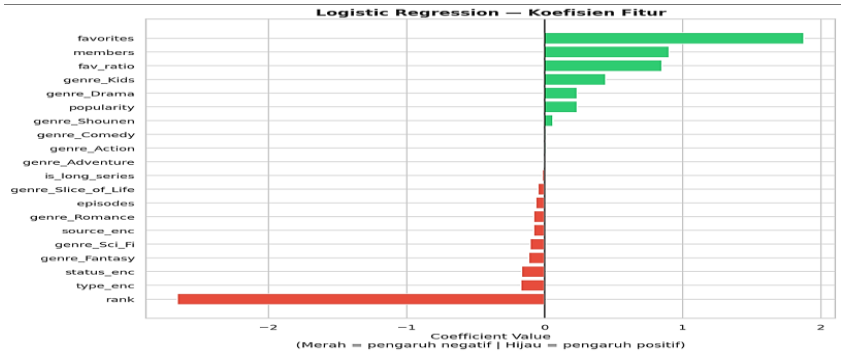
Perbandingan RF vs. *Decision Tree*. Gambar 10 membandingkan *feature importance* kedua model pada union top-10 fitur masing-masing. Random Forest mendistribusikan kepentingan lebih merata antar fitur, sementara Decision Tree memiliki konsentrasi ekstrem pada 1–2 fitur teratas. Hal ini mendukung argumen bahwa Random Forest lebih robust terhadap fitur yang saling berkorelasi (Gupta, S., Adhikari, U., Varshney, S., & Choudhury, 2025).




**Gambar 10.** Perbandingan *Feature Importance*: Random Forest vs. Decision Tree

### *Logistic Regression*

Gambar 11 menampilkan koefisien *Logistic Regression*. Fitur dengan koefisien positif (warna hijau) meningkatkan probabilitas prediksi *high-score*, sedangkan fitur dengan koefisien negatif (warna merah) menurunkannya. Koefisien ini lebih sulit diinterpretasikan secara langsung akibat multikolinearitas antara *members*, *favorites*, dan *rank* (H. Liu, 2025).




**Gambar 11.** Koefisien Fitur — Logistic Regression (Merah = Negatif, Hijau = Positif)

### Cross-Validation

Gambar 12 menampilkan distribusi skor ROC-AUC dari 5-fold stratified *cross-validation*. *Random Forest* memiliki median CV AUC tertinggi dan variasi yang relatif kecil, mengindikasikan stabilitas performa lintas subset data. *Logistic Regression* juga menunjukkan stabilitas baik dengan varians

rendah. *Decision Tree* memiliki median lebih rendah dan variasi yang lebih besar, mengonfirmasi ketidakstabilannya sebagai model tunggal (Yağcı, 2022).

**Implikasi Praktis, Keterbatasan, dan Rekomendasi**

Penelitian ini memperlihatkan bahwa model *Random Forest* berpeluang digunakan sebagai penyaring awal *(*ranker*)* dalam sistem kurasi dan rekomendasi anime, sedangkan dominasi fitur members dan favorites menegaskan perlunya strategi pembangunan komunitas pengguna sejak tahap awal untuk meningkatkan persepsi kualitas anime. Akan tetapi, penelitian ini masih memiliki keterbatasan, antara lain pemakaian sampel sebanyak 300.000 interaksi pengguna, penyederhanaan skor menjadi klasifikasi biner dengan ambang 7,0, kemungkinan multikolinearitas antarfaktor, serta representasi genre yang belum sepenuhnya komprehensif. Penelitian berikutnya disarankan untuk memanfaatkan seluruh dataset, menerapkan time-based split, meneliti algoritma seperti XGBoost dan LightGBM, serta mengintegrasikan analisis *fairness* guna menekan potensi bias model. (Z. Liu et al., 2023).

## SIMPULAN

Penelitian ini membandingkan tiga algoritma klasifikasi *Random Forest*, *Logistic Regression*, dan *Decision Tree* untuk memprediksi *anime high-score* (≥ 7,0) menggunakan dataset *MyAnimeList* dari Kaggle. Pipeline yang dibangun mencakup eksplorasi data, rekayasa fitur multi-tipe (fitur numerik, kategorikal ter-encode, *one-hot genre*, dan fitur turunan), serta evaluasi komprehensif dengan metrik ganda dan *cross-validation*.

Hasil eksperimen secara konsisten menempatkan Random Forest sebagai model terbaik, diikuti *Logistic Regression*, dan *Decision Tree* di posisi terakhir. Keunggulan *Random Forest* terlihat pada seluruh metrik evaluasi maupun kestabilan lintas *fold cross-validation*, yang mencerminkan keunggulan pendekatan ensemble dalam menangani fitur campuran dan ketidakseimbangan kelas. *Logistic Regression* menjadi pilihan yang kompetitif ketika interpretabilitas menjadi prioritas, sementara *Decision Tree* lebih sesuai untuk keperluan eksplorasi aturan sederhana.

Analisis *feature importance* mengungkapkan bahwa fitur keterlibatan pengguna khususnya *members, rank*, dan *favorites* merupakan prediktor paling dominan, jauh melampaui fitur konten seperti genre dan tipe. Temuan ini mengindikasikan bahwa skor komunitas pada MAL lebih mencerminkan eksposur dan keterlibatan pengguna daripada atribut konten semata, yang memiliki implikasi penting bagi desain sistem rekomendasi yang adil dan tidak bias popularitas.

Ke depan, penelitian dapat diperluas dengan menggunakan dataset lengkap, menerapkan time-based validation, mengeksplorasi model *gradient boosting*, dan mengintegrasikan fitur multimodal seperti *embedding* sinopsis untuk meningkatkan representasi konten.

## DAFTAR PUSTAKA

Armenta-Segura, J., & Sidorov, G. (2025). Anime popularity prediction before huge investments: a multimodal approach using deep learning. *PeerJ Computer Science*. https://doi.org/https://doi.org/10.7717/peerj-cs.2715

Chen, J., Dong, H., Wang, X., Feng, F., Wang, M., & He, X. (2020). Bias and Debias in Recommender System: A Survey and Future Directions. *ACM Transactions on Information Systems*.

Gupta, S., Adhikari, U., Varshney, S., & Choudhury, T. (2025). Computational Analysis of Genre Effects on Movie Ratings Using MLP Algorithms. *Journal of Computer Science*. https://doi.org/https://doi.org/10.3844/jcssp. 2025.905.917

Jumaev, G. (2026). Real Time Decision Making and Information Extraction Algorithms Models and Applications. *ACM International Conference Proceedings*. https://doi.org/https://doi.org/10.1145/3726122.3726212

Klimashevskaia, A., Jannach, D., Elahi, M., & Trattner, C. (2024). A survey on popularity bias in recommender systems. *User Modeling and User-Adapted Interaction*. https://doi.org/https://doi.org/10.1007/s11257-024-09406-0

Liu, H. (2025). xploring the Effects of Social Influence on Online Ratings and Reviews. *Proceedings of Business and Economic Studies*. https://doi.org/10.54254/2754-1169/158/2025.19703

Liu, Z., Fang, Y., & Wu, M. (2023). Mitigating Popularity Bias for Users and Items with Fairness-centric Adaptive Recommendation. *ACM Transactions on Information Systems*, *41*(3), 1–27. https://doi.org/10.1145/3564286

Mihara, R. (2018). Involution: a perspective for understanding Japanese animation's domestic business in a global context. *Taylor & Francis on Behalf of the British Association for Japanese Studies Japan Forum*. https://doi.org/10.1080/09555803.2018.1442362

Nachouki, M., Mohamed, E. A., Mehdi, R., & Abou Naaj, M. (2023). Student course grade prediction using the random forest algorithm: Analysis of predictors' importance. *Trends in Neuroscience and Education*, *33*, 100214. https://doi.org/10.1016/j.tine.2023.100214

Nguyen, D., Zigmond, S., Glassco, S., Tran, B., & Giabbanelli, P. J. (2024). ). Big data meets storytelling: using machine learning to predict popular fanfiction. *Social Network Analysis and Mining*. https://doi.org/https://doi.org/10.

1007/s13278-024-01224-x

Pradhan, G., & Tobaben, M. (2024). Hyperparameters in Score-Based Membership Inference Attacks. *Proceedings of Machine Learning Research.*

Rajesh, S., & Kaliaperumal, D. A. (2025). Navigating OTT Choice Overload: Cognitive, Behavioural, and Emotional Effects. *Journal of Consumer Research*. https://doi.org/Navigating OTT Choice Overload: Cognitive, Behavioural, and Emotional Effects

Salman Al-Tameemi, I. K., Feizi-Derakhshi, M.-R., Farhadi, Z., & Feizi-Derakhshi, A.-R. (2025). Hybrid Model for Visual Sentiment Classification Using Content-Based Image Retrieval and Multi-Input Convolutional Neural Network. *International Journal of Intelligent Systems*, *2025*(1). https://doi.org/10.1155/int/5581601

Truong, O., et al. (2024). LeakageDetector 2.0: Analyzing Data Leakage in Jupyter-Driven Machine Learning Pipelines. *MSR 2024 Proceedings*.

Yağcı, M. (2022). Educational data mining: prediction of students' academic performance using machine learning algorithms. *Smart Learning Environments*, *9*(1), 11. https://doi.org/10.1186/s40561-022-00192-z

Zhu, Q., Lo, L. Y., Xia, M., Chen, Z., & Ma, X. (2022). Awareness of Self-Selection Bias in User Ratings. *Proceedings of the ACM on Human-Computer Interaction*. https://doi.org/https://doi.org/10.1145/3555597