

**RANCANG BANGUN WEBSITE REKOMENDASI ANIME BERBASIS
KECERDASAN BUATAN**

TUGAS AKHIR



UNIVERSITAS
MA CHUNG

ALBERT MEIDISON WIJAYA

NIM : 311810002

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNOLOGI DAN DESAIN
UNIVERSITAS MA CHUNG
MALANG
2025**

PERNYATAAN KEASLIAN SKRIPSI

Dengan ini saya menyatakan bahwa isi sebagian maupun keseluruhan Skripsi saya dengan judul “RANCANG BANGUN WEBSITE REKOMENDASI ANIME BERBASIS KECERDASAN BUATAN” adalah benar benar hasil karya intelektual mandiri, diselesaikan tanpa menggunakan bahan-bahan yang tidak diizinkan dan bukan merupakan karya pihak lain yang saya akui sebagai karya sendiri.

Semua referensi yang dikutip maupun dirujuk telah ditulis secara lengkap pada daftar pustaka. Apabila ternyata pernyataan ini tidak benar, saya bersedia menerima sanksi sesuai peraturan yang berlaku.

Malang, 31 Juli 2025



Handwritten signature of Albert Meidison Wijaya.

Albert Meidison Wijaya

NIM. 311810002

**LEMBAR PENGESAHAN
TUGAS AKHIR**

**RANCANG BANGUN WEBSITE REKOMENDASI ANIME BERBASIS
KECERDASAN BUATAN**

Oleh:

ALBERT MEIDISON WIJAYA

NIM : 311810002

dari:

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNOLOGI DAN DESAIN
UNIVERSITAS MA CHUNG**

Telah dinyatakan lulus dalam melaksanakan Tugas Akhir sebagai syarat kelulusan dan berhak mendapatkan gelar Sarjana Komputer (S.Kom.)

Dosen Pembimbing 1,

Dr. Kestrilia Rega Prilianti, M.Si.

NIP. 20120035

Dosen Pembimbing 2,

Prof. Dr.Eng. Romy Budhi, MT.

NIP.20070035

Dekan Fakultas Teknologi dan Desain,



Prof. Dr.Eng. Romy Budhi, ST., MT., M.Pd

NIP. 20070035

RANCANG BANGUN WEBSITE REKOMENDASI ANIME BERBASIS KECERDASAN BUATAN

Albert Meidison Wijaya, Kestrilia Rega, Romy Budhi
Universitas Ma Chung

Abstrak

Anime, sebagai kartun khas Jepang, memiliki keragaman genre dan unsur cerita yang luas, dengan 18 genre dan 322 unsur naratif. Seiring meningkatnya popularitas anime, terutama di kalangan penikmat yang mencari rekomendasi spesifik seperti anime bergenre *iyashikei* untuk mengatasi depresi, dibutuhkan sistem rekomendasi yang mudah diakses dan efektif. Penelitian ini bertujuan mengembangkan *website* rekomendasi anime berbasis kecerdasan buatan yang mudah digunakan dan dapat diakses publik.

Penelitian ini menggunakan algoritma K-Nearest Neighbor (KNN) untuk merekomendasikan anime berdasarkan kemiripan genre dan unsur, dipilih karena akurasi lebih tinggi dibandingkan Naive Bayes setelah pengujian. Analisis sentimen berbasis *tweet* dari Twitter digunakan untuk merekomendasikan anime *iyashikei*, sedangkan rule-based diterapkan untuk *form* filterisasi anime berisi pertanyaan umum penikmat anime ketika mencari anime yang akan ditonton. Data diambil menggunakan API dan dites dengan cross-validation serta penanganan data kosong untuk memastikan keakuratan model.

Website berhasil diimplementasikan tanpa *bug*, dengan waktu perpindahan halaman antara 1–8 detik. Sistem rekomendasi menghasilkan anime dengan akurasi 10-Fold cross-validation sebesar 89,14%, precision 0,87, recall 0,89, F1-Score 0,87, average similarity 0,5733, coverage 0,9411, dan diversity 2,8197. Analisis sentimen *iyashikei* mencapai akurasi 75–98% dengan 10 filter *library dictionary*. Form filterisasi dengan 6 pertanyaan umum berhasil menampilkan rekomendasi yang sesuai. Penelitian ini menyarankan pengembangan fitur personalisasi lebih lanjut untuk meningkatkan pengalaman pengguna.

Kata Kunci: Anime, Website, Sistem Rekomendasi, Kecerdasan Buatan, Analisis Sentimen

DESIGN AND DEVELOPMENT OF AN ARTIFICIAL INTELLIGENCE-BASED ANIME RECOMMENDATION WEBSITE

**Albert Meidison Wijaya, Kestrilia Rega, Romy Budhi
Ma Chung University**

Abstract

Anime, a distinctive form of Japanese animation, features a broad spectrum of 18 genres and 322 narrative elements. With the rising popularity of anime, particularly among fans seeking specific recommendations like iyashikei anime to alleviate depression, there is a demand for an accessible and effective recommendation system. This study aims to develop a user-friendly, publicly accessible anime recommendation website powered by artificial intelligence.

The research employs the K-Nearest Neighbor (KNN) algorithm to recommend anime based on genre and element similarities, selected for its higher accuracy compared to Naive Bayes after testing. Sentiment analysis, utilizing tweets from Twitter, is applied to recommend iyashikei anime, while a rule-based approach is used for an anime filtering form based on common questions from anime enthusiasts. Data is collected via API and processed with cross-validation and handling missing data to ensure model accuracy.

The website was successfully implemented without bugs, with page transition times ranging from 1 to 8 seconds. The recommendation system achieves a 10-Fold cross-validation accuracy of 89.14%, with a precision of 0.87, recall of 0.89, F1-Score of 0.87, average similarity of 0.5733, coverage of 0.9411, and diversity of 2.8197. Sentiment analysis for iyashikei anime yields 75–98% accuracy with 10 library dictionary filters. Filtering form based on 6 questions, delivers accurate recommendations. Further development of personalized features is recommended to enhance user experience.

Keywords: Anime, Website, Recommendation System, Artificial Intelligence, Sentiment Analysis

KATA PENGANTAR

Puji syukur saya panjatkan kepada Tuhan Yang Maha Esa, Yesus Kristus, atas segala rahmat, kasih, dan penyertaan-Nya yang tiada henti, sehingga dapat menyelesaikan laporan Tugas Akhir (TA) dengan judul "**RANCANG BANGUN WEBSITE REKOMENDASI ANIME BERBASIS KECERDASAN BUATAN**", sebagai salah satu syarat untuk menyelesaikan Program Sarjana (S1) Jurusan Teknik Informatika Fakultas Teknologi dan Desain Universitas Ma Chung.

Laporan ini bertujuan untuk menerapkan ilmu yang telah dipelajari selama perkuliahan dan mendapatkan pengalaman serta pengetahuan berharga di bidang koding. Penulis menyadari banyak pihak yang turut membantu sejak awal penyusunan sampai selesainya laporan ini. Pada kesempatan ini, peneliti ingin menyampaikan ucapan terima kasih kepada:

1. Dr.Eng. Romy Budhi, ST., MT., selaku Dekan Fakultas Teknologi dan desain, pembimbing akademik, dosen pembimbing II, sekaligus ketua riset di Ma Chung Human-Machine Interaction Research Center atas bimbingan dan arahnya selama pelaksanaan PKL dan Tugas Akhir.
2. Dr. Kestrilia Rega Prilianti, M.Si., selaku pembimbing utama atas bimbingan, arahan, dan ilmu terutama dalam pengolahan data
3. Seluruh dosen dan staff Ma Chung yang memberikan dukungan doa, ilmu, dan semangat
4. Papa, kakak, dan keluarga besar yang selalu mendukung dan mempercayai meskipun sudah tahun terakhir untuk wisuda dengan waktu yang cukup terbatas
5. Teman di kampus dan luar kampus yang selalu meyakinkan penulis bisa lulus tetap waktu dan memberi semangat

Saya menyadari bahwa laporan ini masih jauh dari sempurna, maka dari itu sangat mengharapkan kritik dan saran yang membangun. Semoga laporan ini dapat memberikan kontribusi bagi pengembangan penelitian di bidang analisis

sentimen, *wesbite*, dan aplikasi teknologi untuk membantu penyembuhan depresi melalui rekomendasi anime.

Malang, 31 Juli 2025



Penyusun

Albert Meidison Wijaya



UNIVERSITAS
MA CHUNG

DAFTAR ISI

KATA PENGANTAR	i
DAFTAR ISI	iii
DAFTAR TABEL	viii
DAFTAR GAMBAR	viii
BAB I	1
PENDAHULUAN	1
1.1 Latar belakang	1
1.2 Identifikasi masalah	3
1.3 Batasan masalah	3
1.4 Rumusan masalah	4
1.5 Tujuan penelitian	4
1.6 Luaran	4
1.7 Manfaat Penelitian	4
1.8 Sistematika Penulisan	5
BAB II	7
TINJAUAN PUSTAKA	7
2.1 Anime	7
2.2 Text Mining	9
2.2.1 Teknik Text Mining	10
2.3 Analisis Sentimen	12
2.4 <i>Application Programming Interface</i> (API)	13
2.4.1 Text Mining Twitter	14
2.4.2 Limit Text Mining Twitter	14
2.4.3 Filter <i>Text Mining</i>	14
2.5 Support Vector Machine (SVM)	16
2.5.1 Metode Evaluasi Model	17
2.5.2 K-Fold <i>cross-validation</i>	18
2.6 <i>Library Dictionary</i>	19
2.6.1 VADER (<i>Valence Aware Dictionary and Sentiment Reasoner</i>)	19
2.6.2 TextBlob	20

2.6.3 SentiWordNet 3.0	20
2.6.4 AFIIN	20
2.6.5 Flair	20
2.7 Rule-based filtering	21
2.8 Sistem Rekomendasi	21
2.8.1 KNN (<i>K Nearest Neighbourhood</i>)	21
2.8.2 Naive Bayes	25
2.9 Flask	26
2.10 ngrok	28
2.11 Studi Terkait	28
2.11.1 Sentimen Analisis Masyarakat Terhadap Manga One Piece di Media Sosial Twitter dengan Metode SVM	28
2.11.2 Sistem Rekomendasi Anime Menggunakan Metode Singular Value Decomposition (SVD) dan Cosine Similarity	29
2.11.3 Penerapan Content-Based Filtering pada Sistem Rekomendasi Anime Berbasis Semantik	29
BAB III	30
ANALISIS DAN PERANCANGAN SISTEM	30
3.1 <i>Flowchart</i> Rancang Bangun Website	30
3.1.1 Alur koding Analisis Sentimen	31
3.1.2 Form Filterisasi Anime	32
3.1.3 Rekomendasi Anime Termirip Berdasarkan Genre dan Unsur	32
3.2 Masalah	33
3.3 Analisis Kebutuhan	33
3.3.1 API	34
3.3.2 Tweet	34
3.3.3 Cara Filterisasi Anime	34
3.4 Studi Literatur	35
3.4.1 Sistem Pencarian	35
3.4.2 API	36
3.5 Desain Sistem	37
3.5.1 Fitur yang dibangun	37

3.5.2 Alur Website	38
3.5.3 <i>Mockup Website</i>	38
3.6 Pengembangan Sistem	46
3.7 Pengambilan Data	46
3.7.1 Pembuatan Excel sebagai database	46
3.7.2 Pembuatan Koding Anime Termirip Berdasarkan Genre dan Unsur	49
3.7.3 Pembuatan Koding Analisis Sentimen	51
3.8 Pembuatan Website Anime	53
3.9 Evaluasi	55
3.9.1 Evaluasi Metode Klasifikasi	55
3.9.2 Evaluasi Website	55
3.10 Perbaikan	56
3.10.1 White-box Testing	56
3.10.2 SVM	57
3.10.3 KNN	57
3.10.4 Naive Bayes	58
3.10.5 Kecepatan	58
3.10.6 Pendapat Pengguna	58
BAB IV	59
HASIL DAN PEMBAHASAN	59
4.1 Hasil Analisis Sentimen	59
4.1.1 Install Node.js	59
4.1.2 Tweet Harvest	60
4.1.3 Fold	61
4.1.4 Hasil Analisis Sentimen Tweet	62
4.1.5 Confusion Matrix	64
4.1.6 <i>Error</i>	68
4.1.7 Wordcloud 1 Kata	69
4.1.8 <i>Wordcloud</i> 2-3 kata (N-gram)	70
4.1.9 Perbandingan Akurasi 9 Sentimen Analisis	71
4.1.10 Excel Analisis Sentimen	76
4.2 Hasil Pengambilan Data	77

4.2.1 Data Anime dari API Anilist	77
4.2.2 Pertanyaan Form Filter Anime	78
4.3 Perbandingan KNN dan Naive Bayes	78
4.4 Pengujian Kodinng	80
4.4.1 <i>Link</i> ngrok	80
4.4.2 <i>Landing page</i>	80
4.4.3 Hasil Pencarian Anime Termirip Berdasarkan Genre dan Unsur ...	84
Blue Lock: EPISODE Nagi	89
4.4.4 Sentimen Analisis Anime Genre <i>Iyashikei</i>	93
4.4.5 Form Filter anime	98
Haikyuu!! 2nd Season	100
4.5 Diskusi	104
BAB V	109
PENUTUP	109
5.1 Kesimpulan	109
5.2 Saran	110
DAFTAR PUSTAKA	111
LAMPIRAN	116
1. Hasil Analisis Sentimen	116
2. Tabel Perbandingan hasil dari 61 judul anime genre <i>iyashikei</i>	122
3. Hasil Pengambilan 5 Data Pertama Anime menggunakan API Anilist	143
4. Hasil akurasi KNN dan Naive Bayes jika anime dipisah <60 jelek, 60-80 bagus, >=80 sangat bagus	146
5. Hasil akurasi KNN dan Naive Bayes jika anime dipisah <65 jelek, 65-80 bagus, >=80 sangat bagus	153
6. Hasil akurasi KNN dan Naive Bayes jika anime dipisah <70 jelek, 70-80 bagus, >=80 sangat bagus	160
7. Koding Sentimen Analisis (Google Colab)	166
8. Koding API Anilist, KNN, dan Pembuatan Website (Google Colab)	167

DAFTAR TABEL

Tabel 4.1	Perbandingan hasil dari 10 judul anime genre <i>iyashikei</i> terpopuler	71
Tabel 4.2	Data Anime <i>Shingeki no Kyojin</i> pada excel anime_data.xlsx	77
Tabel 4.3	Perbandingan Confusion Matrix 6 Model Pencarian Anime Termirip	79
Tabel 4.4	10 Anime termirip dengan <i>Blue Lock</i> berdasarkan KNN Euclidian	89
Tabel 4.5	10 Anime termirip dengan <i>Yuru Camp</i> berdasarkan KNN Euclidian	890
Tabel 4.6	10 Anime termirip dengan <i>Hello World</i> berdasarkan KNN Euclidian	891
Tabel 4.7	10 Anime teratas dengan kriteria rating minimal 80, non-fiksi, cerita asli tamat, tidak ada unsur LGBT+, tidak ada unsur kekerasan, dan cerita tidak berat	100
Tabel 4.8	10 Anime teratas dengan kriteria rating minimal 80, fiksi, cerita asli tamat, tidak ada unsur LGBT+, ada unsur kekerasan, dan cerita berat	101
Tabel 4.9	10 Anime teratas kriteria rating minimal 80, fiksi maupun non fiksi, cerita asli belum tamat, tidak ada unsur LGBT+, tidak ada unsur kekerasan, dan cerita berat maupun ringan	102

DAFTAR GAMBAR

Gambar 2.1 Rekomendasi Myanimelist untuk judul spesifik	8
Gambar 2.2 Rekomendasi Anilist untuk judul spesifik	9
Gambar 2.3 Proses text mining	10
Gambar 2.4 Analisis sentimen terhadap servis	12
Gambar 2.5 Alur Kerja API	13
Gambar 2.6 Hasil <i>search</i> X	15
Gambar 2.7 Ilustrasi Algoritma Support Vector Machine	17
Gambar 2.8 Confusion Matrix	18
Gambar 2.9 Alur kerja K-Fold cross-validation	19
Gambar 2.10 Ilustrasi KNN	22
Gambar 2.11 <i>Euclidian distance</i>	23
Gambar 2.12 Cosine Similirity	24
Gambar 2.13 Jaccard Similarity	25
Gambar 2.14 Naive Bayes	26
Gambar 3.1 <i>Flowchart</i> Koding <i>Wesbite</i>	30
Gambar 3.2 Alur koding Sentimen Analisis	31
Gambar 3.3 Alur Koding <i>Form</i> Kriteria Anime	32
Gambar 3.4 Alur koding rekomendasi anime sejenis	33
Gambar 3.5 <i>Mockup landing page</i>	39
Gambar 3.6 <i>Mockup</i> Rekomendasi Anime Termirip Berdasarkan Algoritma Terbaik antara KNN atau Naive Bayes	41
Gambar 3.7 <i>Mockup</i> Analisis Sentimen Anime Genre <i>Iyashikei</i> <i>Filter library</i>	42
Gambar 3.8 <i>Mockup</i> Perbandingan <i>library</i> Analisis Sentimen Anime Genre <i>Iyashikei</i>	43
Gambar 3.9 <i>Mockup Form</i> Filter Anime	44
Gambar 3.10 <i>Mockup</i> Hasil Filteriasi <i>Form</i>	45
Gambar 4.1 Instalasi Node.js berhasil	59
Gambar 4.2 <i>Harvest No More Tweet Found</i> dan <i>Time Out</i>	60

Gambar 4.3 <i>Harvest</i> Melebihi <i>Limit</i> tapi Berjalan Terus	60
Gambar 4.4 <i>Harvest</i> Melebihi <i>Limit</i>	61
Gambar 4.5 Nilai C, Kernel, Gamma, Fold Terbaik, dan Hasil Tiap <i>fold</i> ...	61
Gambar 4.6 Confusion Matrix VADER <i>Usagi Drop</i>	64
Gambar 4.7 Confusion Matrix TextBlob <i>Usagi Drop</i>	64
Gambar 4.8 Confusion Matrix TextBlob + Pembersihan Data <i>Usagi Drop</i>	65
Gambar 4.9 Confusion Matrix SentiWordNet <i>Usagi Drop</i>	65
Gambar 4.10 Confusion Matrix SentiWordNet + Pembersihan Data <i>Usagi Drop</i>	66
Gambar 4.11 Confusion Matrix AFINN <i>Usagi Drop</i>	66
Gambar 4.12 Confusion Matrix AFINN + Pembersihan Data <i>Usagi Drop</i> ..	67
Gambar 4.13 Confusion Matrix Flair <i>Usagi Drop</i>	67
Gambar 4.14 Confusion Matrix Flair + Pembersihan Data <i>Usagi Drop</i>	68
Gambar 4.15 <i>Error</i> SentiWordNet	68
Gambar 4.16 <i>Error</i> SentiWordNet + Pembersihan Data	69
Gambar 4.17 10 Teratas Kata Positif dan Negatif	69
Gambar 4.18 <i>Worcloud</i> 1 Kata Positif dan Negatif	70
Gambar 4.19 N-Gram Teratas Kata Positif dan Negatif	70
Gambar 4.20 <i>Worcloud</i> 2-3 Kata Positif dan Negatif	71
Gambar 4.21 Excel HasilAkurasi_311810002.xlsx	76
Gambar 4.22 Excel sentimen_analisis.xlsx	76
Gambar 4.23 Perbandingan Akurasi 6 Model Pencarian anime Termirip	79
Gambar 4.24 Halaman <i>About to Visit</i> ngrok	80
Gambar 4.25 <i>Landing page</i>	81
Gambar 4.26 Fitur Cari Anime Termirip Berdasarkan Genre dan Unsur	81
Gambar 4.27 Fitur <i>autocomplete</i>	82
Gambar 4.28 Fitur Analisis Sentimen Anime Genre <i>Iyashikei</i>	83
Gambar 4.29 Penjelasan <i>Form</i> Filter Anime	83
Gambar 4.30 <i>Footer</i>	84
Gambar 4.31 Rekomendasi Anime Termirip dengan Blue Lock (Bagian 1)	85
Gambar 4.32 Rekomendasi Anime Termirip dengan Blue Lock (Bagian 2)	85
Gambar 4.33 Hasil Rekomendasi Jika Tidak Ada <i>Link Streaming</i>	86

Gambar 4.34 <i>Website Streaming Page Not Found</i>	86
Gambar 4.35 <i>Website Streaming jika ada</i>	87
Gambar 4.36 <i>Website Anilist</i>	88
Gambar 4.37 Tombol ke Halaman Awal	88
Gambar 4.38 Halaman Pencarian Anime Termirip Jika <i>Typo</i> Terlalu Parah	93
Gambar 4.39 Perbandingan Anime Terpositif <i>Library VADER</i>	94
Gambar 4.40 Perbandingan Anime Terpositif <i>Library TextBlob</i>	94
Gambar 4.41 Perbandingan Anime Terpositif <i>Library TextBlob Plus</i>	95
Gambar 4.42 Perbandingan Anime Terpositif <i>Library SentiWordNet</i>	95
Gambar 4.43 Perbandingan Anime Terpositif <i>Library SentiWordNet Plus</i>	96
Gambar 4.44 Perbandingan Anime Terpositif <i>Library AFIIN</i>	96
Gambar 4.45 Perbandingan Anime Terpositif <i>Library AFIIN Plus</i>	97
Gambar 4.46 Perbandingan Anime Terpositif <i>Library Flair</i>	97
Gambar 4.47 Perbandingan Anime Terpositif <i>Library Plus</i>	98
Gambar 4.48 Perbandingan 9 <i>Library</i> Terhadap <i>Super Cub</i>	98
Gambar 4.49 Halaman Form Anime dan Peringatan	99
Gambar 4.50 Halaman Filter <i>Form</i> Anime	99
Gambar 4.51 Tombol Edit <i>Form</i> dan Kembali ke Halaman Awal	100
Gambar 4.52 Halaman Filter <i>Form</i> Jika Tidak Ada yang Memenuhi Kriteria	104

UNIVERSITAS
MA CHUNG

BAB I

PENDAHULUAN

1.1 Latar belakang

Di era digital saat ini, media sosial memiliki peran kunci dalam menyebarkan informasi, berdiskusi, dan beropini. Per 6 September 2024 tercatat pengguna sosial media mencapai 5,17 miliar akun atau 63,7% populasi dunia. Rata-rata orang menghabiskan 2 jam 20 menit per hari untuk menggunakan sekitar 7 jenis aplikasi (Backlindo, 2024). Dixon (2024) memberikan data X atau Twitter adalah salah satu media sosial terkenal yang menjadi tempat umum untuk berbagi pandangan, ulasan, berita, dan opini terhadap berbagai topik. Statista (2024) mencatat per Januari 2024, X berada di peringkat 12 dengan 619 juta pengguna aktif dengan 24,69 juta merupakan orang Indonesia. Salah satu keunggulan fitur X dibanding media sosial lain adalah pencarian berdasarkan kesukaan dan *hashtag* terkait tren termasuk bisa mengubah berdasarkan lokasi negara. Dari berbagai jenis tren dan topik yang sering muncul adalah pembahasan anime (Hetler, 2023).

Anime adalah kartun buatan Jepang yang menjadi pembahasan umum di media sosial. Diskusi topik seperti apakah anime bagus atau tidak, dunia produksi, kelanjutan cerita, dan semacamnya. Terdapat dua jenis berdasarkan sumbernya yaitu original dan yang mengadaptasi bahan seperti manga, novel, game, dan *light novel*. Akan tetapi judul anime memiliki kecenderungan sama dengan karya lain contohnya anime *air* dan film bioskop *air*. Namun, jumlah judul anime yang tercatat berjumlah sekitar 26.000 dan terus bertambah di situs MyAnimeList, membuat beberapa orang akan kesusahan dalam memilih judul yang akan ditonton. Faktor yang membuat orang memutuskan menonton selain bagus adalah mengecek adanya genre yang susah diterima seperti LGBT, *incest*, maupun *gore*, cerita yang berhenti di tengah pembuatan, tidak menarik, susah dipahami, dan jumlah episode. Meskipun ada rekomendasi di situs seperti MyAnimeList dan AniList, pembahasan *website* besar dan konten kreator pembahasannya cenderung yang populer sehingga tidak ada patokan bagus tidaknya anime yang kurang dibahas. Hal ini membuat pencarian harus dilakukan manual dan memakan waktu

lama sehingga menjadi masalah terutama untuk orang depresi (Myanimelist, 2024).

Anime memiliki jenis cerita dan beberapa genre cocok untuk bersantai seperti *iyashikei*. Menurut Alodokter (2022), penderita depresi cenderung memiliki gejala seperti tidak berenergi, mudah emosi, bahkan percobaan bunuh diri. Cara meningkatkan efektifitas psikoterapi untuk rehabilitasi depresi adalah melakukan hobi yang positif, menanamkan pemikiran bisa sembuh, dan mengurangi bahkan tidak menggunakan media sosial. Anime yang didesain paling cocok untuk penderita depresi adalah *iyashikei* atau *healing* dan komedi. Jenis cerita yang santai, tidak banyak masalah, dan bahagia. Mudah akses menonton anime dalam berbagai platform membuat penderita bisa menonton di rumah sakit ataupun perawatan mandiri bersama keluarga. Kalau dari judul saja sulit pencarian manual, maka perlu pendekatan lain, yaitu text mining. Dalam kasus ini, *text mining* dapat digunakan untuk menganalisis apakah judul anime dianggap cocok atau tidak untuk penderita depresi.

Text mining adalah proses penggalian informasi yang tersembunyi dalam teks tidak terstruktur. Ini melibatkan teknik komputasi dan statistik untuk menemukan pola dan wawasan. Teks diurai menjadi elemen yang lebih kecil, seperti kata, frasa, atau topik, kemudian menganalisisnya untuk mendapatkan informasi lebih detail tentang isi teks tersebut. Teknik analisis dalam text mining yang akan digunakan adalah analisis sentimen yang berfungsi menentukan teks memiliki sentimen positif atau negatif. Dalam konteks penelitian ini, analisis sentimen dapat digunakan untuk menentukan anime direkomendasikan apabila hasil sentimen mayoritas positif dan sebaliknya (Shiska, 2023).

Pada penelitian ini, akan dikembangkan sebuah media untuk melakukan analisis sentimen untuk menentukan cocok tidaknya anime untuk penderita depresi. Untuk kemudahan agar bisa diakses kapanpun dan dimanapun oleh pengguna, akan dibuat dalam bentuk *website*. Adapun teknik klasifikasi yang digunakan algoritma Support Vector Machine (SVM). SVM merupakan salah satu algoritma klarifikasi dan regresi yang dikhususkan untuk menangani dua kelas sehingga cocok untuk sentimen analisis. Konsep SVM diawali dengan masalah klasifikasi dua kelas sehingga membutuhkan set pelatihan positif dan negatif.

SVM akan berusaha mendapatkan pemisah antar kelas dan memaksimalkan margin (IBM, 2023).

Beberapa hasil penelitian menunjukkan bahwa SVM lebih unggul dibandingkan Naive Bayes dalam pengujian data (Indrayuni, 2018; Salsabila, 2023). Namun, Menurut Mailo dan Lazuardi (2019), Naive Bayes lebih unggul dibandingkan dengan algoritma seperti SVM dan Deep Belief Network dalam pengujian data.

Awalnya, sistem dirancang hanya untuk analisis sentimen menggunakan SVM. Namun, untuk memperluas fungsionalitas dan mendukung kebutuhan pengguna, fitur tambahan berupa form berbasis aturan (rule-based) dan pencarian berdasarkan cerita termirip menggunakan algoritma antara K-Nearest Neighbor (KNN) atau Naive Bayes, yang akan dipilih berdasarkan performa terbaik pada tahap pengujian.

Berdasarkan latar belakang di atas, penulis membuat topik “Rancang Bangun Website Rekomendasi Anime Berbasis Kecerdasan Buatan”.

1.2 Identifikasi masalah

Berdasarkan latar belakang di atas, penulis telah mengidentifikasi permasalahan sebagai berikut:

1. Penderita depresi jika menonton anime yang memiliki unsur kekerasan, *gore*, atau cerita yang terlalu berat, justru meningkatkan sakit depresi.

1.3 Batasan masalah

Batasan masalah yang terdapat pada penelitian ini sebagai berikut:

1. Sumber data sentimen analisis anime genre *iyashikei* hanya berasal dari *tweet* di Twitter berbahasa Inggris.
2. Filter yang digunakan untuk mengambil *tweet* adalah gabungan *keyword* dan *exclude*.
3. Preprocessing *tweet* menggunakan prosedur text mining.
4. Metode yang digunakan untuk pengolahan data analisis sentimen adalah Support Vector Machine.
5. Hasil klasifikasi sentimen analisis berupa sentimen positif dan negatif.

6. Algoritma yang dibandingkan kinerjanya untuk pencarian anime termirip secara genre dan unsur adalah KNN dan Naive Bayes.
7. Algoritma untuk filterisasi form adalah rule-based.
8. Website tidak didesain untuk pengguna *smartphone*.

1.4 Rumusan masalah

1. Bagaimana cara membuat rancang bangun *website* rekomendasi anime berbasis kecerdasan buatan?
2. Bagaimana membandingkan kinerja KNN dan Naive Bayes untuk menghitung kemiripan anime dalam fitur pencarian anime termirip.

1.5 Tujuan penelitian

Tujuan dari penelitian ini adalah:

1. Membuat rancang bangun *website* rekomendasi anime berbasis kecerdasan buatan.
2. Membandingkan kinerja KNN dan Naive Bayes untuk menghitung kemiripan anime dalam fitur pencarian anime termirip.

1.6 Luaran

Luaran dari penelitian ini adalah:

1. Website dengan 3 fitur yaitu rekomendasi anime genre *iyashikei* berdasarkan *tweet* di Twitter, form filterisasi anime, dan rekomendasi anime menggunakan genre dan unsur.
2. Publikasi ilmiah pada jurnal Sainsbertek Fakultas Teknologi dan Desain Universitas Ma Chung

1.7 Manfaat Penelitian

Manfaat dari penelitian ini adalah:

1. Bagi penulis, penelitian ini merupakan eksplorasi teori terhadap analisis sentimen dan pembuatan *website*.
2. Bagi universitas, sebagai tolak ukur pengetahuan mahasiswa.
3. Bagi pembaca, memberikan informasi mengenai sentimen terhadap

judul anime serta referensi penelitian analisis sentimen.

4. Bagi penikmat anime, sebagai tempat mengecek rekomendasi anime genre *iyashikei*, filterisasi anime, dan rekomendasi anime sejenis.
5. Bagi penderita depresi, sebagai aplikasi yang bisa digunakan tahap rehabilitasi.

1.8 Sistematika Penulisan

Berikut adalah urutan sistematika penulisan laporan Tugas Akhir:

- Bab 1: Pendahuluan

Bab ini menjelaskan menjelaskan konteks dan urgensi penelitian dari topik yang diangkat yaitu perlunya *website* rekomendasi anime berbasis kecerdasan buatan untuk memudahkan pengguna menemukan anime yang relevan terutama bagi penderita depresi.

- Bab 2: Landasan Teori

Berisi penjelasan apa itu anime, teori text mining, penggunaan analisis sentimen, penggunaan API untuk pengambilan data secara masif, jenis *library dictionary* untuk perbandingan performa, SVM untuk klasifikasi analisis sentimen, evaluasi model, K-Fold cross-validation untuk mencegah overfitting dan underfitting, rule-based, algoritma KNN dan Naive Bayes untuk pengukuran jarak untuk rekomendasi anime berdasarkan genre dan unsur dengan jarak euclidian, cosine, atau jaccard. Pembuatan *website* dengan flask dan ngrok, serta studi terakit mengenai analisis sentimen dan rekomendasi anime.

- Bab 3: Analisis dan Perancangan Sistem

Menjelaskan langkah-langkah yang diambil untuk mencapai tujuan penelitian. Penggambarkan desain sistem secara keseluruhan untuk 3 fitur yaitu rekomendasi anime genre *iyashikei* menggunakan analisis sentimen, form filterisasi anime, dan rekomendasi anime berdasarkan genre dan unsur. Penjelasan arsitektur *website* dan alur kerja sistem rekomendasi dimulai dari masalah, penentuan kebutuhan user serta pengembang, studi literasi, desain sistem, pengembangan sistem, evaluasi, dan perbaikan. Penggunaan Google Colab, flask, ngrok, penanganan bug, dan pengecekan

koding berjalan dengan sempurna. Hal ini memastikan koding dan *website* dapat berjalan cepat dan akurat agar mendapatkan kepuasan pengguna.

- Bab 4: Hasil dan Pembahasan

Menyajikan hasil implementasi sistem, analisis, dan hasil akhir *website* seperti *landing page* dan hasil tiap fitur. Proses pengumpulan data, filtering, pengecekan hasil secara manual, pembuatan *website*, mulai dari UI hingga integrasi sistem rekomendasi AI. Hasil pengujian sistem, termasuk akurasi rekomendasi dan kecepatan respon disajikan dalam bentuk tabel atau grafik untuk memudahkan pemahaman. Pembahasan mencakup analisis kelebihan dan kekurangan sistem, perbandingan *library dictionary* untuk rekomendasi anime genre *iyashikei*, serta perbandingan pengukuran jarak untuk KNN dan Naive Bayes.

- Bab 5: Kesimpulan dan Saran

Penutup penelitian dengan menyimpulkan temuan-temuan utama. *Website* rekomendasi anime berbasis AI berhasil dibangun dan memenuhi tujuan penelitian. Kesimpulan mencakup pencapaian teknis, manfaat yang dirasakan pengguna, serta kemampuan tiap algoritma dan kombinasi terbaik. Saran diberikan untuk pengembangan lebih lanjut seperti penambahan fitur seperti rekomendasi berdasarkan mood, filter genre spesifik, atau penggunaan dataset yang lebih besar dan beragam. Saran juga mencakup penggunaan algoritma lain dan kecepatan pemrosesan data dan perpindahan halaman *website*.

BAB II

TINJAUAN PUSTAKA

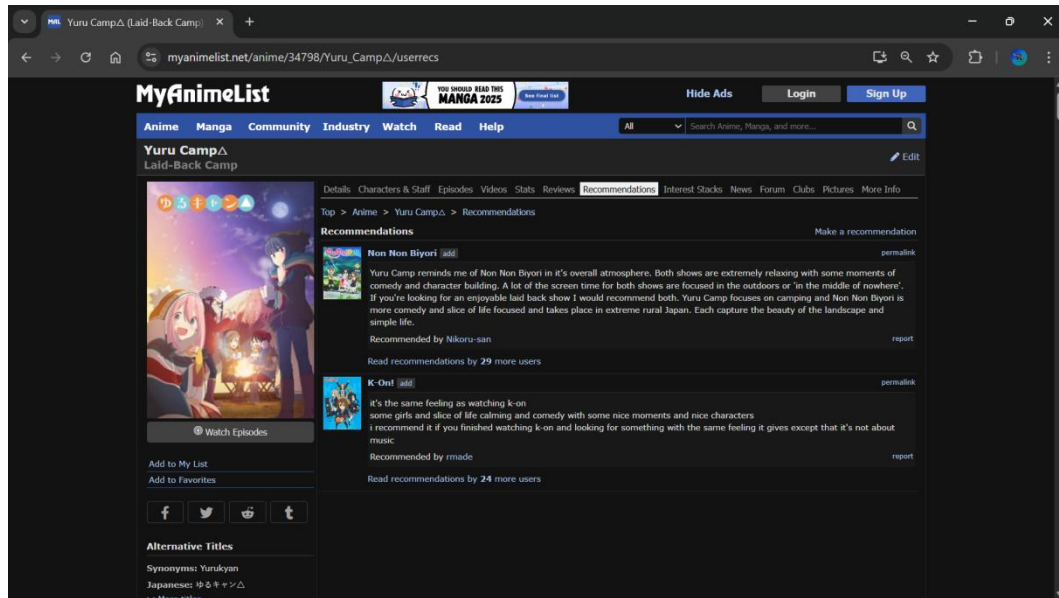
2.1 Anime

Berdasarkan *website* Myanimelist (2025), anime adalah kartun buatan Jepang dengan jenis cerita yang tidak terbatas. Setiap judul anime dibuat oleh 1 studio atau lebih dengan mayoritas judul berjumlah 12 episode x 24 menit atau movie sekitar 2 jam. Anime dapat digolongkan menjadi 5 pasar, yaitu *kids* untuk anak-anak, *shounen* untuk remaja cowok, *seinen* untuk pria dewasa, *shoujou* untuk remaja cewek, dan *josei* untuk perempuan dewasa. Tiap *website* memiliki pandangan yang berbeda antar genre. Selain itu beberapa ada yang menggunakan tema atau unsur sebagai penjelasan. Pada *website* Anilist, terdapat 18 genre seperti *action*, *comedy*, *romance*, dan sebagainya. Untuk unsur cerita dalam Anilist sebanyak 332 jenis merupakan penjelasan detail dari genre seperti sifat karakter, lokasi, jenis penceritaan semisal *disability* tentang karakter disabilitas, *work* tentang pekerjaan, *iyashikei* merupakan cerita yang tenang dan minim masalah, dan sebagainya.

Myanimelist dan Anilist sudah mencantumkan hal di atas termasuk tempat untuk menonton anime, mayoritas di Crunchyroll, Bilibili, Iqiyi, Youtube, Netflix, dan yang lain tergantung pemasaran pihak studio. Penikmat anime lama biasanya mengecek informasi pada kedua *website* tersebut sebelum menonton untuk memperhatikan detail dan *review* penonton lain sebelum memutuskan menonton. Penikmat baru biasanya menonton yang sedang populer di media sosial seperti X, Instagram, Youtube, atau perbincangan di lingkungan sekitar. Permasalahan yang sering dihadapi penikmat baru adalah kurangnya pemahaman jika membaca detail *website* atau jika ingin mendapatkan detail di media sosial, terkadang mereka malu menanyakan di kolom komentar atau menanyakan langsung ke pemilik postingan. Setelah mereka menemukan anime yang bagus dan cocok, ketika mencari anime sejenis, mereka akan kebingungan dengan pilihan yang banyak dan tanpa adanya fasilitas instan akan membuat mereka kehabisan waktu.

Myanimelist merupakan *website* anime terpopuler secara global. Secara isi, Myanimelist menjelaskan secara detail mengenai anime, studio, pengisi suara, dan

lain-lain. Rekomendasi yang diberikan Myanimelist dilakukan dengan cara mencari anime yang ditonton, klik tab *recommendations*, maka akan muncul rekomendasi judul, 1 komen utama, dan tambahan komen pengguna lain.



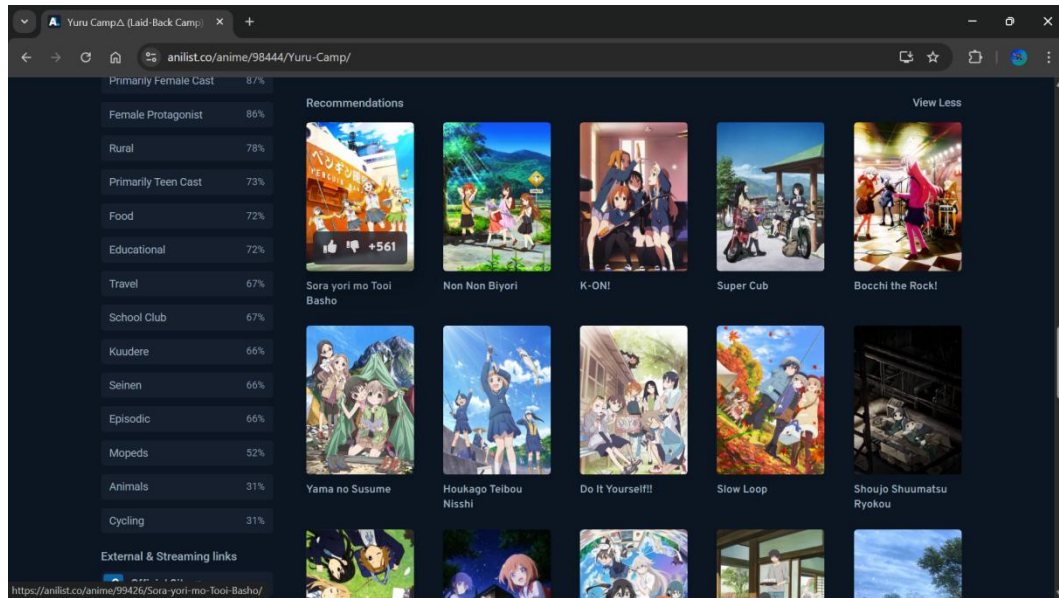
Gambar 2.1 Rekomendasi Myanimelist untuk judul spesifik

Sumber: Myanimelist.net

Pada gambar 2.1, jika penonton menyukai *Yuru Camp*, anime yang direkomendasikan adalah *Non Non Biyori*, *K-on!*, dan seterusnya. Sistem rekomendasi dari pengguna dengan pemberian *review* dan alasan kecocokan sangat membantu, akan tetapi anime yang kurang terkenal atau baru tayang akan mendapatkan rekomendasi yang sedikit sehingga kebanyakan sistem rekomendasi oleh pengguna kebanyakan akan merekomendasikan yang terkenal saja. Selain itu penulisan rekomendasi yang panjang mungkin membuat beberapa orang tidak nyaman karena harus membaca tulisan yang panjang.

Anilist merupakan website anime yang lebih detail terkait anime terutama pada unsur yang terdapat pada anime. Sistem rekomendasi Anilist dilihat pada halaman judul spesifik, turun ke bawah hingga ada tulisan *recommendations*. Rekomendasi hanya memberikan judul serta foto anime yang direkomendasikan dan selisih *like* dengan *dislike* sehingga bisa minus. Tidak diketahui alasan kenapa

bisa anime direkomendasikan untuk anime lain apakah dipilih manual oleh pengembang atau dengan perhitungan matematis.



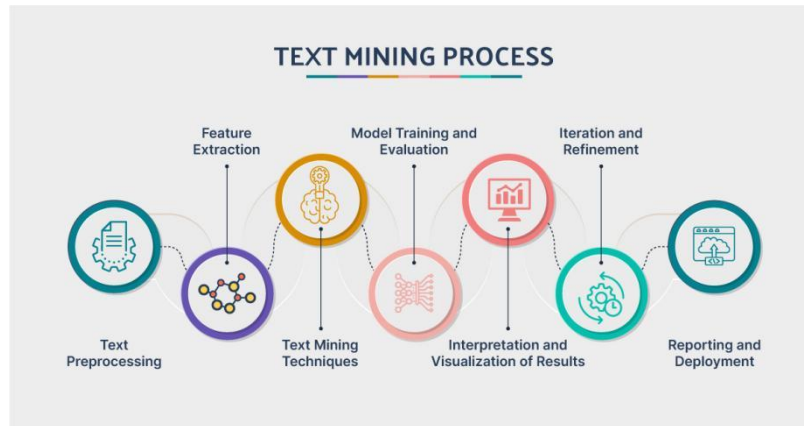
Gambar 2.2 Rekomendasi Anilist untuk judul spesifik

Sumber: Anilist.co

Pada gambar 2.2, jika kalian suka *Yuru Camp*, maka *Sora yori mo Tooi Basho* disarankan minimal oleh 561 pengguna karena tidak diketahui jumlah pasti *like* dan *dislike*. Selain itu, disarankan anime *Non Non Biyori*, *K-on*, *Super Cub*, dan seterusnya. Jumlah rekomendasi Anilist lebih banyak dibandingkan Myanimelist akan tetapi karena tidak ada informasi munculnya rekomendasi, Anilist kurang dijadikan patokan rekomendasi oleh beberapa penikmat anime.

2.2 Text Mining

Text Mining adalah analisis teks yang berfokus pada penggalian informasi yang bermakna dari data teks yang tidak terstruktur seperti postingan media sosial, email, ulasan pelanggan, dan artikel berita karena tidak memiliki struktur yang jelas dengan hasil akhir berupa informasi berharga (Shiksha Online, 2024).



Gambar 2.3 Proses text mining

Sumber : Shiska.com

Pada gambar 2.3, text mining dimulai dengan preprocessing dan pengambilan fitur yang akan digunakan. Teknik-teknik text mining akan dijalankan kemudian dimasukkan ke dalam model dan dievaluasi. Agar hasil mudah dipahami pengguna sebaiknya ditampilkan secara visual. Agar mencegah overfitting dan underfitting, dilakukan iterasi seperti *cross-validation* kemudian hasil akhirnya dilaporkan dan dijalankan pada sistem.

2.2.1 Teknik Text Mining

Text Mining memiliki beberapa teknik dan memiliki tujuan masing-masing. Tidak setiap *library* memakai semua fungsi. Berikut beberapa tekniknya:

1. Text Preprocessing

Text preprocessing adalah transformasi data teks mentah ke format yang sesuai untuk analisis. Hal ini mencakup menghapus tanda baca, mengubah teks menjadi huruf kecil, tokenisasi (membagi teks menjadi kata atau token), menghapus kata yang sering muncul tanpa makna seperti *the* dan *is*, dan *stemming* (mereduksi kata menjadi bentuk dasarnya). Setiap kata akan dicocokkan dan diberikan skor seperti *library* AFIN yang dibuat oleh Nielsen (2011). Misalnya, "abandon" dan turunannya seperti "abandoned" serta "abandons" diberi skor -2, sedangkan "absolve" dan turunannya seperti "absolved" serta "absolves" diberi skor 2, dengan rentang skor dari -5 (sangat negatif) hingga +5 (sangat positif).

2. Text Classification

Text Classification memberikan kategori atau label yang telah ditentukan pada dokumen teks berdasarkan isinya. Algoritma seperti Naive Bayes, Support Vector Machines (SVM), Convolutional Neural Networks (CNN) atau Recurrent Neural Networks (RNN) biasanya digunakan untuk klasifikasi teks.

3. Analisis Sentimen

Analisis sentimen digunakan untuk menentukan sentimen atau opini yang diungkapkan dalam sebuah teks menjadi positif, negatif, atau netral menggunakan metode rule-based, algoritma Machine Learning, dan model Deep Learning.

4. *Named Entity Recognition* (NER)

NER bertujuan untuk mengidentifikasi dan mengklasifikasikan entitas bernama, seperti nama orang, lokasi, organisasi, dan tanggal di dalam data teks. NER melibatkan penggunaan teknik seperti pencocokan pola, metode berbasis aturan, atau algoritma Machine Learning untuk mengenali dan mengekstrak.

5. Pemodelan topik

Pemodelan topik adalah teknik yang digunakan untuk menemukan topik dalam kumpulan dokumen. Pemodelan topik membantu memahami konsep utama yang dibahas dalam data teks. Algoritma untuk pemodelan topik termasuk Latent Dirichlet Allocation (LDA) dan Non-Negative Matrix Factorization (NMF).

6. Pengelompokan Teks

Pengelompokan dokumen yang serupa berdasarkan isinya. Hal ini membantu mengidentifikasi pola dan hubungan dalam koleksi besar dokumen teks. Teknik seperti k-means, hierarchical clustering, dan density-based clustering dapat digunakan untuk pengelompokan teks.

7. Peringkasan Teks

Teknik peringkasan teks bertujuan untuk menghasilkan ringkasan dari teks panjang secara ekstraktif atau abstraktif. Peringkasan ekstraktif melibatkan

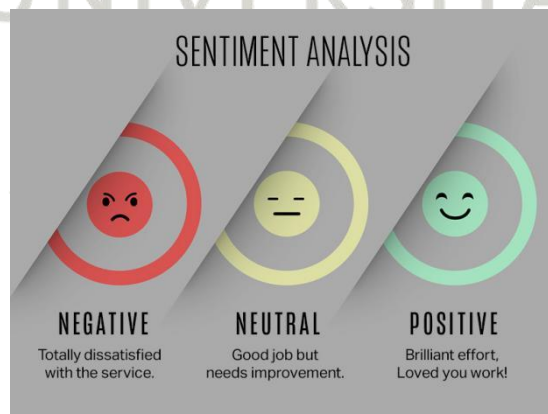
pemilihan kalimat atau frasa penting dari teks asli, sedangkan peringkasan abstraktif melibatkan pembuatan kalimat baru untuk menangkap informasi.

2.3 Analisis Sentimen

Analisis sentimen media sosial adalah proses mengumpulkan dan memahami persepsi pengguna terhadap produk, layanan, atau merek menggunakan algoritma klasifikasi dan *natural language processing* (NLP) untuk mengevaluasi emosi di balik interaksi media sosial. Dengan menganalisis *like*, komentar, *share*, dan *mention* dapat memperoleh wawasan berharga mengenai faktor pendorong emosional yang memengaruhi keputusan pembelian. Analisis sentimen media sosial memberi latar belakang percakapan pelanggan dan menunjukkan apakah merek mendapatkan perhatian yang baik atau buruk. Dan jika demikian, mengapa? (Lown, 2024).

Sentimen Analisis memiliki 4 langkah yaitu :

1. Melihat *mention* pada sosial media.
2. Memisahkan menjadi positif, negatif, bahkan netral.
3. Gunakan analisis sentimen untuk sosial media.
4. Menerapkan analisis sentimen sosial pada strategi meningkatkan daya ingat merek dan beresonansi lebih baik dengan target.



Gambar 2.4 Analisis sentimen terhadap servis

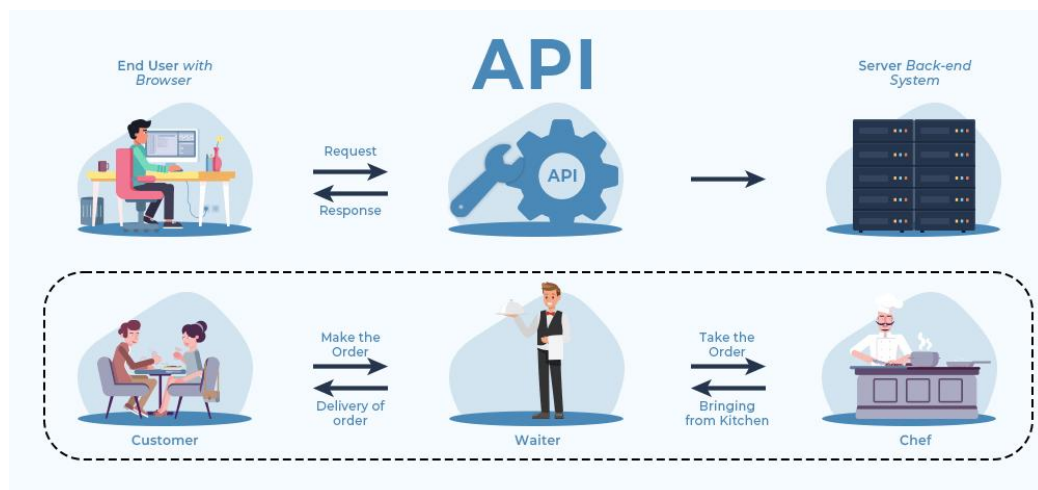
Sumber: imerit.net

Pada gambar 2.4, kalimat sangat tidak puas dengan servis diklasifikasikan sebagai negatif atau secara emosi marah. Kerja bagus tetapi butuh peningkatan

dianggap netral atau secara emosi biasa saja, sementara usaha luar biasa, kerja bagus dikategorikan sebagai positif atau secara emosi tersenyum. Analisis sentimen hanya mengklasifikasikan kalimat ke dalam kelas umumnya positif, netral, dan negatif terpisah dengan analisis emosi teks.

2.4 Application Programming Interface (API)

Goodwin (2024) menyatakan bahwa API menyederhanakan pengembangan aplikasi dan perangkat lunak dengan memungkinkan pengembang untuk mengintegrasikan data, layanan, dan kemampuan dari aplikasi lain daripada membuat dari awal. Pemilik aplikasi juga dapat berbagi data dan fungsi kepada berupa paket data kecil yang relevan dengan permintaan tertentu dari pihak ketiga. API memungkinkan berbagi hanya informasi yang diperlukan untuk menjaga detail sistem internal tetap tersembunyi. Contoh penggunaan API adalah X yang dapat digunakan untuk mengambil dan menganalisis data X secara terprogram serta membangun percakapan di X. Anilist, *website* anime menggunakan API sehingga pengguna dapat mengambil data terkait anime dengan total judul sekitar 21.000 dan akan terus bertambah.



Gambar 2.5 Alur Kerja API

Sumber: IBM.com

Pada gambar 2.5, dijelaskan cara kerja API. Pengguna akan meminta data pada *website* seperti *tweet* pada Twitter atau spesifikasi data anime pada Anilist.

API akan mengecek apakah data boleh diberikan kepada pengguna dengan respon *error* atau data tersebut diberikan.

2.4.1 Text Mining Twitter

Ahmed (2024) melakukan text mining pada beberapa *platform* dan menemukan bahwa X memiliki beberapa keunggulan, antara lain:

1. Twitter memiliki banyak topik seperti kejadian luar biasa, berita, dan *event*.
2. Twitter mempermudah mencari dan mengikuti percakapan dengan fitur *search* dan muncul pada pencarian google.
3. Twitter memiliki *hashtag* yang memudahkan pengumpulan, penyortiran, dan perluasan pencarian.
4. API Twitter lebih terbuka dan dapat diakses dibandingkan dengan media sosial lainnya. Hal ini meningkatkan ketersediaan bagi peneliti.
5. Tweet memiliki panjang yang konsisten (≤ 140). Sementara itu, Facebook postingannya lebih panjang dan bukan gambar seperti Instagram.
6. Sifat data Twitter yang *real-time*.
7. Analisis jaringan digunakan untuk memvisualisasikan hubungan antara orang untuk lebih memahami struktur percakapan.

2.4.2 Limit Text Mining Twitter

Twitter membuat peraturan limit *post*, *following*, dan jumlah *post* yang bisa dibaca oleh pengguna termasuk text mining dengan batas kecepatan 900 permintaan per 15 menit. Jika melebihi akan muncul pesan { "errors": [{ "code": 88, "message": "Rate limit exceeded" }] } (X.Corp, 2024).

2.4.3 Filter Text Mining

Satria (2024) menunjukkan jika data yang diambil untuk text mining dari X dapat dispesifikan agar *tweet* yang didapat sangat detail. Beberapa filter yang dapat dilakukan adalah:

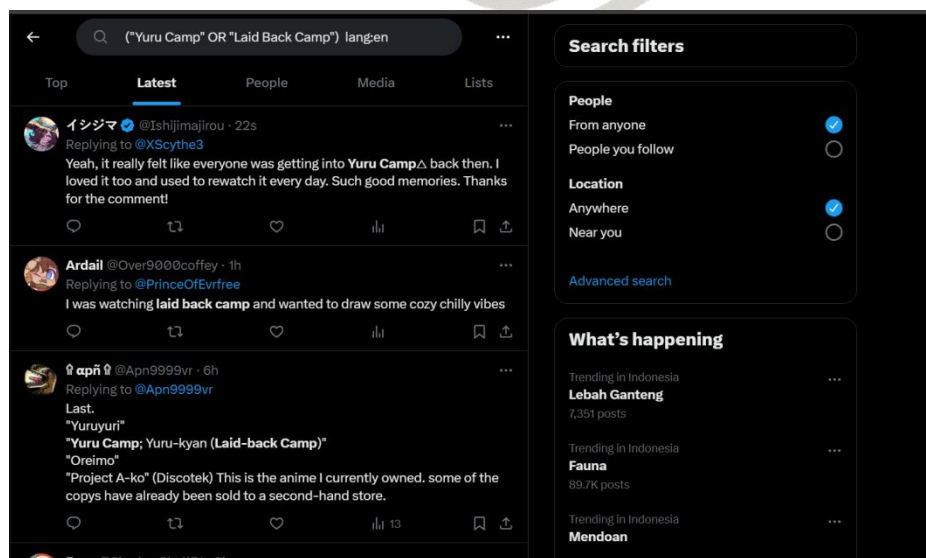
1. `search_keyword` : beberapa kata dan *hashtag*.

2. limit = maksimal data yang akan diambil.
3. lang = bahasa yang disesuaikan dengan *library dictionary*.
4. since: = rentang waktu pertama.
5. until: = rentang waktu terakhir.
6. from:username = akun yang melakukan *post*.
7. - keyword = pengecualian *keyword*.
8. --tab = *Latest* untuk *tweet* terbaru atau *Top* untuk *tweet* terpopuler
9. twitter_auth_token = token unik tiap akun untuk akses ke API twitter.

```

1. filename = `(Keyword1).csv`
2. search_keyword = (Keyword1) (Hastag1) since:2023-04-01
   until:2024-04-01 lang:id'
3. limit = 1000
4.
5. !npx -y tweet-harvest@2.6.1 -o "{filename}" -s
   "{search_keyword}" --tab "LATEST" -l {limit} --token
   {twitter_auth_token}

```



Gambar 2.6 Hasil *search X*

Sumber: X.com

Gambar 2.6 menunjukkan hasil *advanced twitter* dengan OR dengan limitasi bahasa Inggris dengan menampilkan *tweet* berisi *Yuru Camp*, *Laid Back*

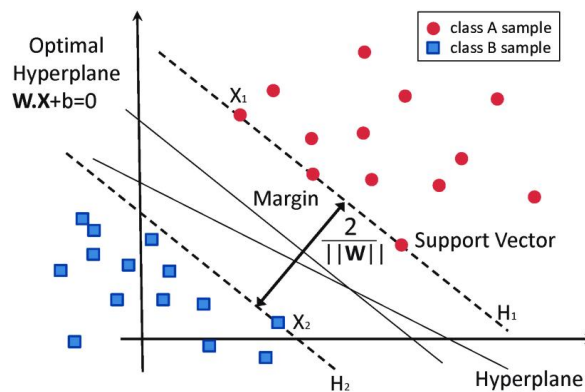
Camp, atau keduanya. Penggunaan fungsi AND dan OR akan memaksimalkan spesifikasi agar data tidak tercampur dan jumlahnya banyak.

2.5 Support Vector Machine (SVM)

Plaosan (2024) menyatakan SVM adalah algoritma pembelajaran mesin yang diawasi untuk mengklasifikasikan data dengan menemukan garis optimal atau *hyperplane* yang memaksimalkan jarak antara setiap kelas dalam ruang N dimensi ke dalam dua kelas. Garis-garis yang bersebelahan dengan *hyperplane* optimal dikenal sebagai vektor pendukung karena melewati titik-titik data yang menentukan margin maksimal. Hal ini memungkinkan untuk menggeneralisasi dengan baik pada data baru dan membuat prediksi klasifikasi yang akurat.

SVM digunakan dalam banyak hal pemrosesan bahasa alami seperti analisis sentimen, deteksi *spam*, dan pemodelan topik. SVM cocok untuk data-data ini karena berkinerja baik dengan data berdimensi tinggi. Cara Kerja SVM sebagai berikut:

1. Membagi data dalam *train* dan *test*.
2. Buat dan evaluasi model impor modul SVM. Latih data latih pada pengklasifikasi dan prediksi respons dengan membandingkan akurasi set pengujian dengan nilai prediksi. Dapat menggunakan metrik evaluasi seperti skor *accuracy*, *f1-score*, *precision*, dan *recall*.
3. Penyetelan *hyperparameter* untuk meningkatkan kinerja model SVM menggunakan pencarian *grid* dan *cross-validation* yang akan melakukan iterasi melalui kernel, regularisasi (C), dan nilai gamma yang berbeda untuk menemukan kombinasi terbaik.



Gambar 2.7 Ilustrasi Algoritma Support Vector Machine

Sumber: Zaid Al-Sawaff

Gambar 2.7 menunjukkan kelas A dan B terpisah dengan baik oleh *hyperplane* lurus $W.X+b=0$.

2.5.1 Metode Evaluasi Model

Dalam Machine Learning, precision dan recall adalah dua metrik yang paling penting dalam menentukan keakuratan model. Recall adalah kemampuan model untuk menemukan semua kasus yang relevan dalam kumpulan data, sedangkan precision adalah kemampuannya untuk mengidentifikasi hanya titik data yang relevan. Keduanya memainkan peran penting dalam klasifikasi yang tidak seimbang, yang melibatkan distribusi data ketika terlalu banyak titik data yang masuk ke dalam satu kelas.

$$\text{Precision} = \text{True Positives} / (\text{True Positives} + \text{False Positives})$$

$$\text{Recall} = \text{True Positives} / (\text{True Positives} + \text{False Negatives})$$

Untuk menemukan perpaduan yang optimal antara precision dan recall, kita dapat menggabungkan kedua metrik tersebut dengan menggunakan F1-score.

$$\text{F1-score} = 2 ((\text{Precision} * \text{Recall}) / (\text{Precision} + \text{Recall}))$$

Untuk klasifikasi biner menunjukkan empat hasil yang berbeda: true positive, false positive, true negative, and false negative. Nilai aktual membentuk kolom dan nilai prediksi membentuk baris (Koehrsen & Urwin, 2024).

		Actual Values	
		1 (Positive)	0 (Negative)
Predicted Values	1 (Positive)	TP (True Positive)	FP (False Positive) <i>Type I Error</i>
	0 (Negative)	FN (False Negative) <i>Type II Error</i>	TN (True Negative)

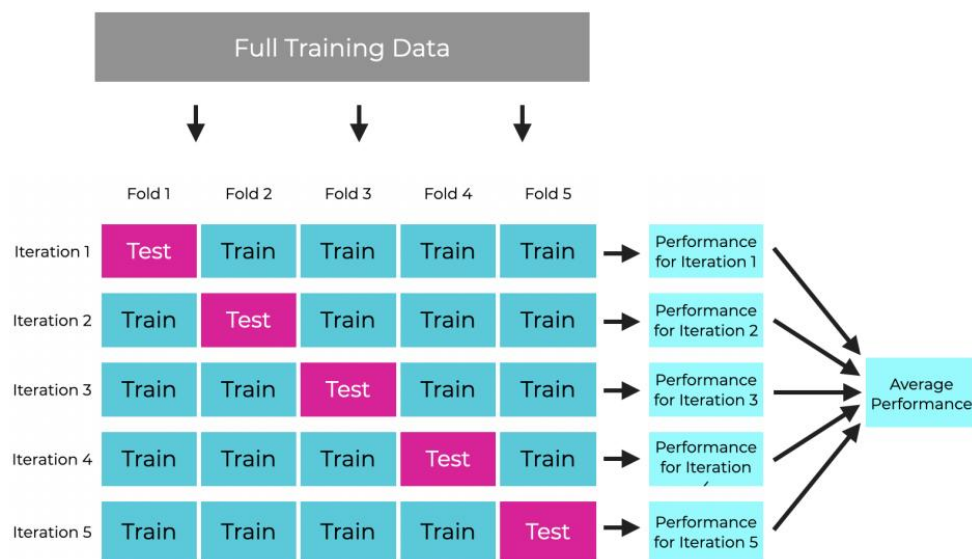
Gambar 2.8 Confusion Matrix

Sumber: Kaggle.com

2.5.2 K-Fold *cross-validation*

Cross-validation berfungsi untuk mencegah overfitting dan underfitting ketika model dilatih terlalu baik pada data pelatihan dan berkinerja buruk pada data baru yang tidak terlihat dan sebaliknya. Dengan mengevaluasi model pada beberapa set validasi, Cross-validation memberikan estimasi yang lebih realistis tentang kinerja generalisasi model, yaitu kemampuannya untuk berkinerja baik pada data baru yang tidak terlihat. Keunggulan lainnya adalah penggunaan data lebih efisien karena setiap observasi digunakan untuk pelatihan dan pengujian.

Cross-validation terbaik adalah K-Fold *Cross-validation*. Dataset dibagi menjadi k subset atau folds, kemudian melakukan pelatihan pada semua subset tetapi menyisakan satu subset (k-1) untuk evaluasi model yang telah dilatih. Iterasi dijalankan sebanyak k kali dengan subset yang berbeda yang dicadangkan untuk tujuan pengujian setiap kali (Sharma, 2024).



Gambar 2.9 Alur kerja K-Fold cross-validation

Sumber: sharpsightlabs.com

Pada gambar 2.9, data dibagi menjadi 5 sama besar atau 80% *train* dan 20% *test*, fold 1 berfungsi sebagai *test* saat iterasi 1 dan menjadi *train* saat iterasi 2. Hal ini terus dijalankan hingga fold terakhir menjadi data *test*. Hasil tiap fold kemudian akan dirata-rata.

2.6 Library Dictionary

Library berisi kamus dan peraturan penulisan bahasa memegang peran penting dalam analisis sentimen. Kelengkapan kata, ketepatan dalam menentukan kata bermakna ganda, penanganan singkatan, dan lain-lain. Berikut 5 *Library* yang akan digunakan dalam penelitian ini :

2.6.1 VADER (*Valence Aware Dictionary and Sentiment Reasoner*)

VADER adalah sebuah modul Python nltk.sentiment yang secara khusus dibuat untuk bekerja dengan teks dalam media sosial, termasuk bahasa dari konteks lain. VADER mampu mendeteksi seberapa positif atau negatif sentimen dari sebuah teks ketika data yang sedang dianalisis tidak diberi label. VADER menggunakan leksikon kata-kata yang berhubungan dengan sentimen untuk menentukan sentimen keseluruhan (Malde, 2020).

2.6.2 TextBlob

TextBlob adalah library Python yang menggunakan *Natural Language Toolkit* (NLTK). Library ini mendukung analisis teks dan memberikan informasi tentang polaritas dan subjektivitas kalimat. Polaritas berkisar antara $[-1,1]$, dengan -1 berarti negatif dan 1 berarti positif. Subjektivitas berkisar antara $[0,1]$, yang menunjukkan seberapa banyak opini pribadi terdapat dalam teks. Misalnya, kalimat "Saya tidak suka contoh ini" mendapatkan polaritas -1 dan subjektivitas 1, sementara kalimat lainnya bisa menghasilkan nilai 0.0 jika polaritasnya seimbang (Shah, 2020).

2.6.3 SentiWordNet 3.0

SentiWordNet 3.0 adalah sumber daya leksikal yang dirancang untuk mendukung klasifikasi sentimen dan aplikasi penggalian opini. Versi ini merupakan pengembangan dari SentiWordNet 1.0, dihasilkan melalui proses anotasi otomatis terhadap semua *synset* WordNet berdasarkan tingkat positif, negatif, dan netralitasnya. Algoritma yang digunakan untuk anotasi otomatis, yang mencakup langkah *random-walk* untuk memperbaiki skor, selain dari langkah pembelajaran *semi-supervised* yang sebelumnya digunakan (Baccianella, Esuli, & Sebastiani, 2010).

2.6.4 AFINN

AFINN adalah leksikon kata-kata bahasa Inggris yang diberi nilai valensi bilangan bulat antara minus lima dan plus lima. Kata-kata tersebut telah diberi label secara manual oleh Finn Årup Nielsen pada tahun 2009-2011 (Perry, 2011).

2.6.5 Flair

Flair adalah *library* NLP dari Zalando Research yang berfokus pada analisis sentimen dengan pendekatan Deep learning. Menggunakan pelabelan urutan, Flair dapat menangkap sentimen setiap kata dalam teks. *Library* ini menyediakan model pra-latih dalam berbagai bahasa dan memungkinkan *fine-tuning* dengan data spesifik. Keunggulannya terletak pada ekstraksi informasi

kontekstual, yang membantu memahami sentimen dalam teks kompleks (Tiwari, 2023).

2.7 Rule-based filtering

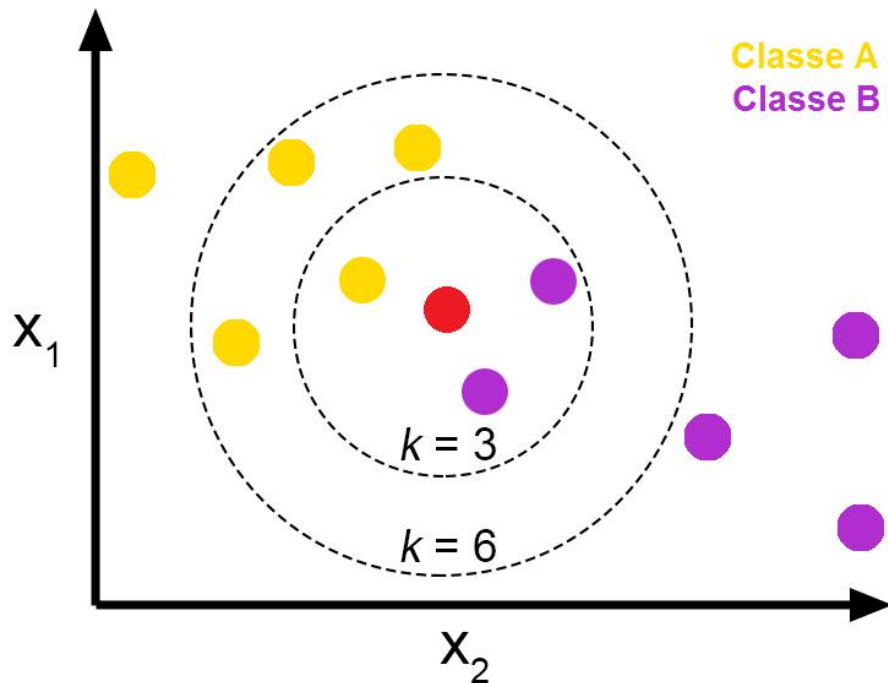
Dataset disiapkan dengan berbagai kriteria atau fitur yang akan difilter, penyiapan database bisa dengan mengambil data menggunakan API secara langsung dan dapat diproses dengan difilter dengan besaran atau aturan tertentu. Sistem akan mengecek isi dari input pertanyaan pada formulir menggunakan fungsi *if*, *then*, *and*, dan *or* dengan isi dataset. Hasil akan keluar berdasarkan syarat yang sesuai, apabila tidak ada yang memenuhi maka tidak ada hasil. Pertanyaan yang dibuat dipastikan relevan, faktual, sudah berkonsultasi dengan ahli untuk menentukan pola, dan memiliki semua opsi jawaban sedetail mungkin agar pengguna tidak kebingungan dalam menjawab (Engati, 2024).

2.8 Sistem Rekomendasi

Memberi rekomendasi untuk pengguna merupakan fitur umum dalam *website* terutama ketika pilihan yang ditawarkan dalam jumlah banyak. Setidaknya tercatat 21.345 judul anime yang telah tayang, sedang tayang, dan akan tayang pada Anilist. Oleh karena itu, untuk memudahkan penikmat anime dalam menentukan anime yang ditonton selanjutnya semirip mungkin, maka semua genre dan unsur akan diambil dan dicari anime dengan tingkat kemiripan tertinggi. Algoritma KNN dan Naive Bayes merupakan pilihan umum dalam memberikan rekomendasi

2.8.1 KNN (*K Nearest Neighbourhood*)

Algoritma KNN adalah metode *supervised* Machine Learning sederhana untuk klasifikasi dan regresi untuk memprediksi kedekatan antar titik data dan mengambil sejumlah k data terdekat. Namun, kinerjanya dipengaruhi oleh pilihan k dan metrik jarak, sehingga penyetelan parameter menentukan hasil optimal.



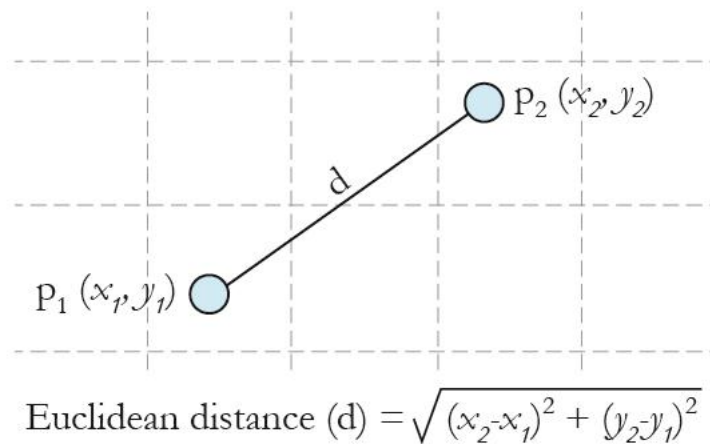
Gambar 2.10 Ilustrasi KNN

Sumber: Medium.com

Cara kerja KNN sebagai berikut:

1. Muat data.
 2. Inisialisasi nilai k.
 3. Untuk mendapatkan kelas yang diprediksi, lakukan iterasi dari 1 hingga jumlah total titik data pelatihan. Hitung jarak antara data uji dan setiap baris set data pelatihan menggunakan jarak euclidean sebagai metrik Jarak data kategorikal.
 4. Urutkan jarak yang dihitung dalam urutan menaik berdasarkan nilai jarak. Dapatkan k baris teratas dari larik yang diurutkan. Dapatkan kelas yang paling sering muncul dari baris-baris ini. Kembalikan kelas yang diprediksi. KNN dapat menggunakan beberapa jenis cara menghitung jarak antara 2 titik seperti euclidian, cosine, atau jaccard, ketiganya cocok untuk jenis data teks.
1. Euclidian Distance
Euclidian Distance digunakan untuk mencari jarak antara dua vektor bernilai riil untuk menghitung jarak nilai riil dengan menormalkan

data ke dalam skala yang seimbang agar hasilnya baik.



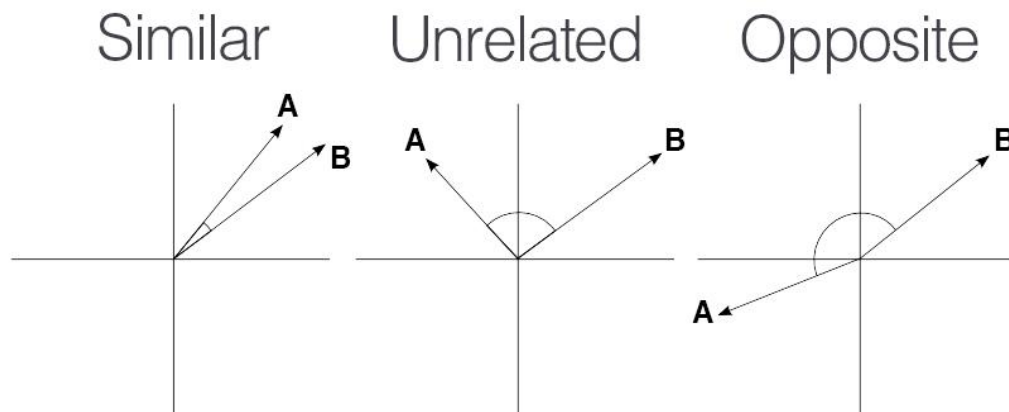
Gambar 2.11 *Euclidian distance*

Sumber: analyticsvidhya.com

Pada gambar 2.11, Data akan ditaruh dalam 2 dimensi x dan y lalu dihitung dengan rumus pitagoras.

2. Cosine Similarity

Cosine similarity adalah metrik matematis yang digunakan untuk mengukur kemiripan antara dua vektor dalam ruang multi-dimensi dengan menghitung cosinus sudut. Hal ini sangat berguna dalam analisis teks, di mana teks diubah menjadi vektor. Setiap dimensi vektor mewakili kata dari dokumen, dengan nilainya yang menunjukkan tingkat kepentingan kata tersebut.



Gambar 2.12 Cosine Similarity

Sumber : dev.to

Pada gambar 2.12, semakin dekat jarak 2 vektor akan mendekati 0 derajat dan saat bernilai berkebalikan menyentuh 180 derajat. Cos 0 derajat bernilai 1, cos 90 derajat bernilai 0, cos 180 derajat bernilai -1. Hasil cosine similarity memiliki rentang -1 hingga 1, berikut rumus perhitungan *cosine similarity*:

$$Sc(x, y) = x \cdot y / \|x\| \times \|y\|$$

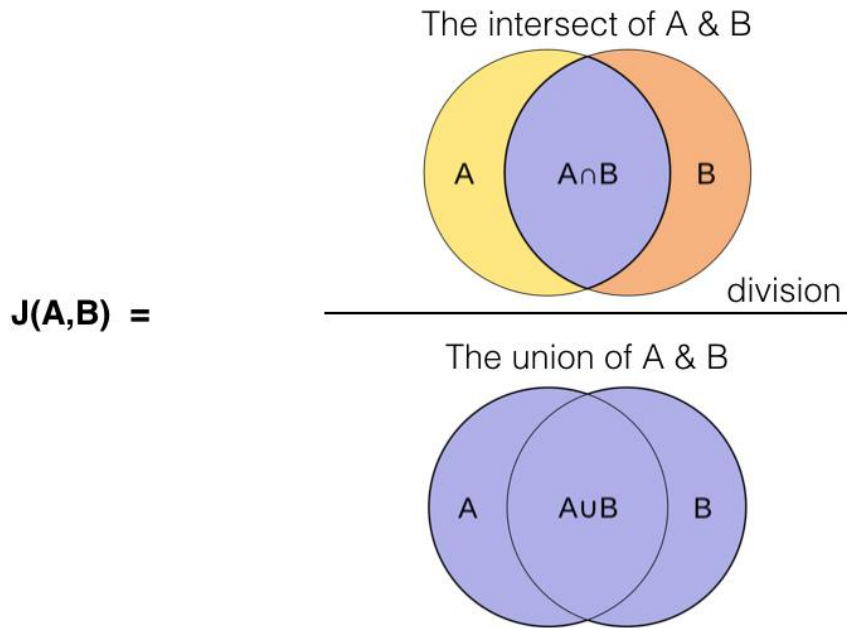
$x \cdot y$ = hasil kali dari vektor 'x' dan 'y'.

$\|x\|$ dan $\|y\|$ = panjang dari dua vektor 'x' dan 'y'.

$\|x\| \times \|y\|$ = hasil kali dari dua vektor 'x' dan 'y'.

3. Jaccard Similarity

Jaccard similarity adalah sebuah statistik untuk mengukur kemiripan antara dua set data. Ini diukur sebagai ukuran perpotongan dua set dibagi dengan ukuran gabungannya. Nilainya berada di antara 0 dan 1. di mana 0 berarti tidak ada kemiripan sama sekali dan mendekati 1 berarti kemiripan mutlak.



Gambar 2.13 Jaccard Similarity

Sumber: Kaggle.com

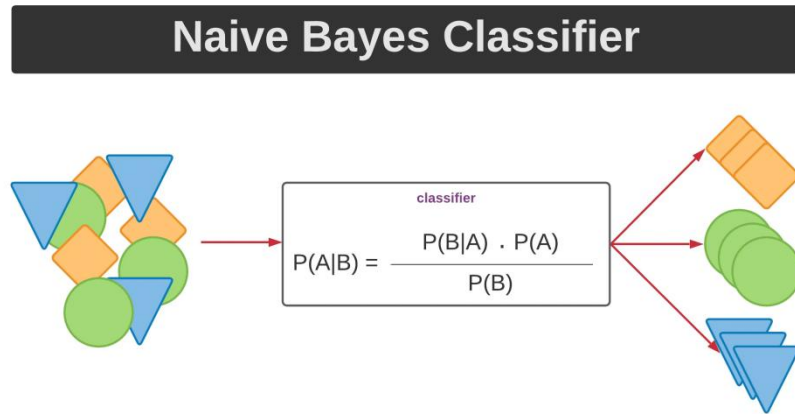
Jaccard hanya menghitung jumlah perpotongan antar data tanpa mempertimbangkan nilai data, sehingga jika ada data beririsan, data pertama bernilai 100% dan data kedua bernilai 1% akan bernilai 1 karena beririsan. Dari ketiga jenis jarak, euclidian lebih cocok untuk perhitungan kemiripan anime menggunakan fitur genre, unsur dengan persentase untuk KNN dan cosine similarity untuk Naive Bayes. Fitur yang dihitung seperti kriteria anime dengan jawaban ya, tidak, dan apa saja akan dibahas lebih detail pada bab 3.

2.8.2 Naive Bayes

Naive Bayes adalah algoritma klasifikasi berbasis probabilitas yang menerapkan teorema Bayes dengan asumsi independensi antar fitur. Algoritma ini menghitung probabilitas kelas berdasarkan fitur, lalu memilih kelas dengan probabilitas tertinggi. Naive Bayes cocok untuk klasifikasi teks dan data berdimensi tinggi karena sederhana dan efisien.

Jenis Naive Bayes ada 3, Gaussian Naive Bayes untuk data kontinu dengan distribusi normal. Multinomial Naive Bayes untuk data diskrit seperti frekuensi kata atau fitur kategori. Bernoulli Naive Bayes untuk data biner (0/1).

Data anime merupakan fitur kategorikal seperti genre dan unsur anime. Algoritma ini mengasumsikan fitur independen dan menghitung probabilitas kelas berdasarkan frekuensi kemunculan fitur (Scikit-learn, 2025).



Gambar 2.14 Naive Bayes
Sumber : Pemrograman Matlab

Cara kerja Multinomial Naive Bayes sebagai berikut (GeeksforGeeks, 2025):

1. Prior Probabilty: Hitung probabilitas awal setiap kelas (misal, $P(\text{kelas} = \text{"direkomendasikan"})$) dari data pelatihan.
2. Likelihood Probabilty: Hitung probabilitas fitur kategorikal muncul pada kelas tertentu menggunakan distribusi multinomial.
3. Klasifikasi: Gabungkan prior dan likelihood untuk memprediksi kelas dengan probabilitas tertinggi.

2.9 Flask

Flask adalah kerangka *web* Python yang kecil dan ringan dengan fleksibilitas tinggi yang lebih mudah diakses oleh para pengembang baru hanya dengan menggunakan satu *file* Python. Flask dapat diperluas dan tidak memaksakan struktur direktori tertentu. Flask menggunakan Jinja *template engine* untuk membuat halaman HTML secara dinamis menggunakan konsep umum Python seperti variabel, perulangan, daftar, dan sebagainya. Website jika

diperlukan menggunakan SQLite. Berikut tahapan membuat *website* menggunakan Flask di Google Colab:

1) Insatlasi Flask

```
!pip intall Flask
```

2) Membuat dasar aplikasi

```
1. from flask import Flask
2.
3. app = Flask(__name__)
4.
5. @app.route('/')
6. def hello():
7.     return 'Hello, World!'
```

@app.route adalah sebuah dekorator yang mengubah fungsi Python biasa menjadi fungsi tampilan Flask, yang mengubah nilai kembalian fungsi menjadi respons HTTP untuk ditampilkan oleh klien HTTP. Input '/' setelah @app.route() untuk menandakan bahwa fungsi ini akan merespons permintaan web untuk URL /, yang merupakan URL utama. Fungsi view hello() mengembalikan string 'Hello, World!' sebagai respon. Contoh tampilan *output*:

Output

```
* Serving Flask app "hello" (lazy loading)
* Environment: development
* Debug mode: on
* Running on http://127.0.0.1:5000/ (Press CTRL+C to quit)
* Restarting with stat
* Debugger is active!
* Debugger PIN: 813-894-335
```

3) HTML

HTML berfungsi menampilkan informasi yang dapat ditampilkan di *web browser*. Flask menyediakan fungsi pembantu `render_template()` agar pengelolaan HTML lebih mudah dengan menulis kode HTML dalam

berkas .html atau bisa langsung mengkode di dalam `render_template()`. Untuk memperbaiki tampilan setiap bagian dapat ditambahkan CSS.

4) Membuat Database

Jika *website* membutuhkan penyimpanan data dalam jumlah besar maka dibutuhkan database untuk mengambil sebagian data menggunakan filter, Database semisal SQLite harus dikoneksikan dengan koding *website*.

5) Tahap Lanjutan

Website dengan 4 tahap di atas sudah cukup untuk dijalankan namun dapat ditambahkan beberapa hal yaitu:

1. Menampilkan daftar postingan di halaman awal.
2. Menampilkan satu postingan lengkap di halaman terpisah berdasarkan ID.
3. Menghasilkan *error* 404 jika postingan tidak ditemukan.
4. Fungsi tingkat lanjut pada *website* semisal CRUD (*Create, Read, Update, Delete*).

2.10 ngrok

Ngrok adalah alat yang digunakan untuk membuat *tunnel* aman dari internet ke *localhost*, memungkinkan server lokal seperti *web* server atau API diakses secara publik melalui URL sementara. Tujuan penggunaan ngrok adalah memfasilitasi pengujian, *debugging*, atau demonstrasi aplikasi yang berjalan di lingkungan lokal tanpa perlu *deploy* ke server eksternal agar dapat diakses oleh pihak lain melalui URL yang dihasilkan, sehingga memudahkan proses evaluasi dan kolaborasi (Ngrok, 2025)

2.11 Studi Terkait

2.11.1 Sentimen Analisis Masyarakat Terhadap Manga One Piece di Media Sosial Twitter dengan Metode SVM

Penelitian ini diharapkan dapat memberikan tanggapan masyarakat yang mendalam tentang keberhasilan *One Piece* dalam berinteraksi dengan penggemarnya. Hasil analisis sentimen masyarakat terhadap manga One Piece di

media sosial Twitter dengan metode SVM yang didapat sejumlah 93,1% sentimen positif dan 6,9% sentimen negatif. Tingkat akurasi yang didapatkan untuk sentimen yaitu negatif 0,50 untuk precisionnya, 0,06 untuk recallnya, 0,11 untuk f1-scorenya dan 17 untuk support. Tingkat akurasi yang didapatkan untuk sentimen positif yaitu 0,94 untuk precision, 1,00 untuk recall, f1-score 0,97 dan 259 untuk support (Hrikesaa, Pratamaa, & Raharjaa, 2020).

2.11.2 Sistem Rekomendasi Anime Menggunakan Metode Singular Value Decomposition (SVD) dan Cosine Similarity

Penelitian ini bertujuan mencari anime dengan kemiripan tertinggi menggunakan genre dengan algoritma cosine similarity dan metode SVD. SVD digunakan untuk menganalisis, mengurangi dimensi, dan merekonstruksi data bertujuan mengurai matriks menjadi tiga matriks yang lebih sederhana. Hasil dari penelitian ini mengularkan anime yang paling mirip *One Piece* adalah *One Piece movie* setelah dibandingkan dengan 12.292 judul di website Myanimelist.

2.11.3 Penerapan Content-Based Filtering pada Sistem Rekomendasi Anime Berbasis Semantik

Tujuan dari penelitian ini adalah untuk membantu penikmat anime dalam merekomendasikan anime dengan menggunakan metode Content-Based Filtering, merekomendasikan suatu dokumen berdasarkan jumlah kemunculan kata yang ada pada sisi pengguna kemudian dilakukan pembobotan akhir dari rumus TF-IDF hingga menghasilkan sebuah rekomendasi. Kriteria yang digunakan untuk menentukan sebuah rekomendasi adalah judul, tipe, studio, genre, tahun rilis, dan pengisi suara karakter. Sistem Rekomendasi Anime Berbasis Semantik yang dapat memperkaya informasi yang diberikan sehingga menjadi lebih baik dalam pendefinisianannya agar memungkinkan komputer dapat memahami informasi. Penentuan rekomendasi tergantung pada judul anime yang disukai oleh pengguna dan dengan TF-IDF menampilkan rekomendasi anime, jika belum ada anime yang disukai maka website tampil secara default tanpa adanya rekomendasi.

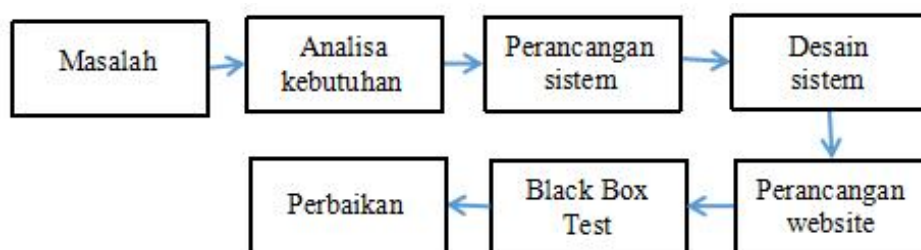
BAB III

ANALISIS DAN PERANCANGAN SISTEM

3.1 *Flowchart Rancang Bangun Website*

Website rekomendasi anime berbasis kecerdasan buatan perlu dispesifikasikan fitur apa saja yang perlu dibuat agar dapat menjawab masalah penikmat anime dalam mencari rekomendasi anime terutama bagi penderita depresi agar tidak salah menonton anime tertentu yang bisa memperparah penyakitnya. Analisis kebutuhan dibutuhkan supaya apa yang diinginkan pengguna tapi bisa dibuat peneliti sehingga harus mendapatkan data yang diperlukan untuk membuat program. Program bisa saja mengalami *update* sehingga diperlukan studi literatur seperti penggunaan algoritma, API, dan pembuatan *website*.

Ketika data bisa dikodingkan sesuai keinginan pengguna, aplikasi perlu dibuat desain UI agar pengguna tidak kebingungan dalam mengetahui dan menggunakan fitur. Ketika desain *website* sudah dibuat, maka dimulai pengembangan sistem, seperti tempat *coding* seperti google colab, akses *localhost* dengan flask, pembuatan website, dan lainnya. Jika sistem sudah dibuat, maka perlu evaluasi seperti hasil confusion matrix, hasil yang sesuai, pengecekan *bug* aplikasi, apakah data mengeluarkan hasil dengan cepat, dan menanyakan *beta tester* apakah UI dan hasil sudah bagus. Jika terdapat *error*, maka perlu ada perbaikan dan penyesuaian lebih lanjut, pencarian *bug*, *hyperparameter tuning*, bahkan mengganti algoritma. Penjelasan di atas terangkas pada flowchart dibawah:



Gambar 3.1 *Flowchart Koding Website*

3.1.1 Alur koding Analisis Sentimen

Analisis kebutuhan dilakukan dengan riset melalui Twitter dan dapat disimpulkan sebagai berikut:

1. Kalimat *tweet* memiliki beberapa *input* unik seperti *emoji* dan *emoticon*
2. Penggunaan *keyword* yang umum akan menyebabkan data tercampur dengan penelusuran yang lain sehingga perlu spesifikasi *keyword*.

Perancangan dimulai dari pemilihan bahasa pemrograman, *library*, algoritma yang dibutuhkan, dan alur koding. Rancangan dibuat sesederhana, seringan, dan secepat mungkin tanpa mengurangi fitur yang diperlukan bertujuan mendapatkan statistik yang bagus. Berikut alur kerja koding sentimen analisis:



Gambar 3.2 Alur koding Sentimen Analisis

Pada gambar 3.2, Alur sentimen analisis dimulai dengan mengambil *tweet* dengan kata spesifik tanpa duplikat. Setiap *tweet* akan digolongkan menjadi positif dan negatif agar bisa diprediksi menggunakan Support Vector Classification (SVC). Hasil aktual dan prediksi akan dimunculkan dalam confusion matrix dan dilakukan analisa *error*.

K-Fold cross-validation perlu dijalankan agar tidak mengalami overfitting dan underfitting. Persentase suka akan diperkuat dengan *wordcloud* 1 kata dan *N-gram* (2,3) sehingga dapat menambahkan faktor positif dan negatif setiap judul. Hasil akhir dari perhitungan seluruh nilai token akan dicek menggunakan SVM.

Hasil sentimen akan diberi label positif jika nilai akhir ≥ 0 dan negatif jika < 0 . Data ini akan diproses dengan SVM di mana SVM akan mempelajari pola kalimat baru akan menjadi positif atau tidak. Perlu pengaturan *hyperparameter* seperti apakah berbentuk linear atau non linear, TF-IDF vectorizer, confusion matrix, dan menggunakan k-fold *cross-validation* untuk memastikan semua jenis data dalam *test* dan *train* bisa diprediksi.

3.1.2 Form Filterisasi Anime

Setiap judul anime yang diambil menggunakan API anilist akan diambil data judul, rating, genre dan unsur, lalu dijalankan dengan alur sebagai berikut:



Gambar 3.3 Alur Koding *Form* Kriteria Anime

Pada gambar 3.3, semua judul akan diambil tanpa terkecuali. Data yang diambil akan diterapkan ke rule-based. Dalam pembangunan *website*, data dijadikan pada database semisal SQLite namun dengan ukuran file kecil cukup dibuat dalam excel dan diupload ke google colab dan pastikan *file* dibaca dan bisa dikeluarkan pada koding. Form pertanyaan dibuat sesuai kebutuhan pengguna, simpan pilihan pengguna, kemudian dicocokkan pada database. Setiap hasil yang cocok akan ditampilkan pada *website* dengan urutan rating tertinggi.

3.1.3 Rekomendasi Anime Termirip Berdasarkan Genre dan Unsur

Mengambil seluruh genre dan unsur, kemudian genre jika ada bernilai 1 dan 0 jika tidak ada. Unsur jika ada, diambil angka tanpa persen dan bernilai 0 jika tidak ada. Enam pertanyaan pada rule-based akan berubah dari ya menjadi 1

dan tidak menjadi 0. Data tersebut dimasukan pada k-nn dan dicek kedekatannya. Berikut alur kerja koding rekomendasi anime sejenis:



Gambar 3.4 Alur koding rekomendasi anime sejenis

Pada gambar 3.4, data yang sudah diubah ke numerik akan dicari kedekatannya ke dalam KNN dengan jarak euclidian dan ditampilkan ke *website*.

3.2 Masalah

Berdasarkan postingan dan komentar di sosial media, terdapat beberapa masalah tentang kurangnya fitur pada *website* yang sudah ada yaitu:

1. Anime genre *iyashikei* yang cocok untuk penderita depresi tidak banyak yang diketahui.
2. Banyak pengguna media sosial menginginkan anime dengan cerita mirip dengan yang diposting orang lain, namun terkadang tidak ada respon.
3. Banyak orang yang menyesal menonton suatu anime terutama yang memiliki unsur kekerasan berlebih, cerita dengan unsur LGBT, atau cerita yang terlalu berat.

3.3 Analisis Kebutuhan

Kebutuhan antara pengguna dan pengembang berbeda. Kebutuhan pengguna adalah sebagai berikut:

1. Aplikasi yang selalu *update*, lengkap, dan proses pencarian cepat.
2. Akses yang mudah kapanpun dan di manapun.
3. Pembuatan akun bersifat opsional.
4. UI yang mudah dan nyaman agar fitur dapat digunakan dengan tepat.
5. Pengisian pertanyaan jika dibutuhkan sedikit dan tidak bertele-tele.

Sementara kebutuhan pengembah baru diketahui setelah mengetahui apa yang pengguna inginkan yaitu sebagai berikut:

3.3.1 API

Data terupdate sangat cepat bahkan tiap hari bisa ada penambahan judul dan informasi. Mengingat banyaknya judul, maka pembuatan secara manual berserta perhitungan akan membutuhkan waktu lama. Oleh karena itu, mengambil data pada *website* lain bisa dilakukan selama *website* tersebut isinya benar, terupdate, dan memiliki API. API adalah cara mengambil data secara masif dengan melakukan permintaan kepada *website* dengan menuliskan format sesuai *website* tersebut. Jika *website* tidak memiliki API, maka data tidak bisa diambil atau jika menggunakan database seperti kaggle, data tersebut sudah kadaluarsa.

3.3.2 Tweet

Untuk mendapatkan rekomendasi terupdate, maka pencarian bisa dilakukan dengan mengambil postingan terbaru di media sosial. Dari banyaknya media sosial, X lebih mudah karena data berbentuk teks, memiliki API, serta bisa melakukan filter *tweet* sesuai keinginan kita sedetail mungkin. Kumpulan *tweet* yang sudah terdownload dalam excel wajib dicek, apakah memang sudah sesuai keinginan kita seperti tidak tercampur dengan pembahasan lain. Tweet akan dicari sentimennya apakah positif atau negatif, dibuat modelnya dengan *library dictionary* dan algoritma SVC. Jika model memiliki akurasi tinggi dengan semua judul, maka hasil akhir perhitungan akan dikeluarkan persentase positif dengan perhitungan $(\text{True Positive} + \text{False Negative}) / \text{Total Tweet}$ yang akan ditampilkan kepada pengguna.

3.3.3 Cara Filterisasi Anime

Pertanyaan pada *form* berisi pertanyaan umum terkait anime yang akan ditonton. Pertanyaan umumnya memiliki 3 jawaban ya, tidak, dan apa saja (terserah) sehingga dengan rule-based sederhana, mencocokkan input *form* dengan database akan mengeluarkan hasil yang sesuai. Namun, perlu dibuat pertanyaan sesuai kebutuhan penikmat anime yaitu:

1. Rating anime dari 1-10.
2. Apakah sumber asli (*manga*, *light novel*, atau *novel*) sudah tamat?
3. Cerita fiksi atau non-fiksi.
4. Apakah ada unsur *lesbian*, *gay*, *bisexual*, dan *transgender* (LGBT+)?
5. Apakah ada unsur kekerasan berlebihan?
6. Apakah ceritanya berat?
7. Apakah anime sesuai dengan karya asli?
8. Bagaimana nuansa cerita sedih, tenang, bergejolak, atau bahagia?
9. Apakah anime sudah mencakup semua material sumber asli?
10. Apakah grafis anime bagus?

Pertanyaan ini harus dicek apakah bisa difilter melalui API, misalnya cerita fiksi atau non-fiksi bisa difilter menggunakan genre dan unsur. Pengecekan grafis bagus tidak bisa dilakukan karena tidak ada data nilai grafis pada *website*, termasuk nuansa cerita, apakah anime sudah dibuat sampai halaman terakhir karya asli, dan apakah ada perubahan besar dari karya asli. Oleh karena itu, pertanyaan 7 sampai 10 tidak bisa dikodingkan dan tidak akan ada pada *website*.

3.4 Studi Literatur

Dalam pembuatan *website*, akan diperlukan beberapa tambahan terkait pengambilan data dan pembuatan *website*

3.4.1 Sistem Pencarian

Google dan AI seperti deepseek adalah contoh aplikasi pencarian umum sementara *website* seperti MyAnimeList dan Anilist dikhususkan untuk anime. Namun, semuanya memiliki kesamaan, perlunya memasukan *input*, menekan tombol spesifik atau *keyboard enter* dan mengeluarkan 2 jenis hasil, data yang sesuai atau pemberitahuan jika tidak ada yang sesuai. Penambahan seperti rekomendasi pilihan di bagian bawah yang sejenis sangat memudahkan terutama saat tulisan *input* terdapat kesalahan penulisan. Sistem pencarian apabila tidak bisa dilakukan dengan memasukan *input*, maka perlu dibuatkan *form* berisi

beberapa pertanyaan. Jawaban yang disediakan harus mencakup semua kemungkinan tanpa terkecuali. Semua *input* akan dicocokkan dengan database sehingga perlunya pembuatan database.

3.4.2 API

Pengambilan data secara masif hanya bisa dilakukan dengan API, maka perlu mengetahui *website* yang memiliki fitur API. Setelah mengetahui kebutuhan pengguna, maka kita dapat mengambil hanya data yang dibutuhkan agar *website* yang dibangun cepat dan tidak terlalu berat. Pengambilan data harus mengetahui struktur HTML sederhana seperti nama kelas dan perpindahan halaman *website*. Contoh API pada Anilist untuk mencari anime yang berhubungan:

```
1.query ($id: Int!) {  
2.  Media (id: $id) {  
3.    id  
4.    title {  
5.      romaji  
6.    }  
7.    relations {  
8.      edges {  
9.        relationType  
10.       node {  
11.         id  
12.         title {  
13.           romaji  
14.         }  
15.       }  
16.     }  
17.   }  
18. }  
19.}
```

Koding di atas mencari judul berdasarkan id, di dalam id terdapat judul dan hanya mengambil romaji (abjad Romawi untuk bahasa Jepang), relasi (hubungan dengan judul lain semisal lanjutan) berbentuk *node*. Pemahaman perpindahan halaman dibutuhkan karena perlunya pengambilan data berulang.

3.5 Desain Sistem

Website perlu dirancang apa saja fitur yang akan dibuat, penataan UI, tombol, dan perpindahan halaman.

3.5.1 Fitur yang dibangun

Website memiliki 3 fitur utama yang akan dibangun yaitu:

1. Ranking anime genre iyashikei di Twitter

Berisi penyerderhanaan hasil PKL “Analisis Sentimen pada Twitter terhadap Rekomendasi Judul Anime Genre Iyashikei Menggunakan Metode Support Vector Machine”. Rekomendasi berdasarkan sentimen analisis *tweet* mengeluarkan hasil akhir antara lain judul anime, total *tweet*, persentase positif yang dihitung berdasarkan $(\text{True Positive} + \text{False Negative}) / \text{Total Tweet}$, *link streaming*, dan *link* Anilist.

2. Rekomendasi anime berdasarkan kemiripan genre dan unsur pada Anilist

Pengguna mengisi judul untuk mendapatkan maksimal 50 anime termirip yang diurutkan dari persentase kemiripan tertinggi berdasarkan genre dan unsur di Anilist menggunakan algoritma terbaik antara KNN dan Naive Bayes yang akan dijelaskan lebih lanjut di bab 4. Rekomendasi akan memunculkan judul, rating, genre, 3 unsur dengan persentase tertinggi, *link streaming*, dan *link* Anilist. Hal ini memudahkan pengguna mencari anime dengan cerita termirip dengan anime yang sudah ditonton dan bisa menonton langsung di *link streaming* yang disediakan.

3. Form filterisasi anime

Formulir berisi 6 pertanyaan, rating diisi dengan angka 1-100. Apakah karya asli sudah tamat, fiksi, unsur LGBT+, kekerasan ekstrim, dan apakah cerita berat memilih opsi antara ya, tidak, dan apa saja. Apa saja berarti jawaban ya dan tidak keduanya terpilih. Hasil rekomendasi akan berisi judul, rating, genre, 3 unsur dengan persentase tertinggi, *link streaming*, dan *link* Anilist dan diurutkan berdasarkan rating tertinggi.

3.5.2 Alur Website

Website memiliki alur perpindahan halaman dari *landing page* sesuai fungsi yang dipakai seperti navigasi, tombol *submit*, atau mengklik *link*, berikut 3 fungsi perpindahan halaman pada rancangan website:

1. Pengguna memasukan judul pada *seacrh box* pada fitur rekomendasi berdasarkan genre dan unsur dan mengklik tombol cari, akan berganti ke halaman yang menampilkan judul, rating, genre, tingkat kemiripan, *link streaming*, dan *link* Anilist kemudian memunculkan hingga 50 hasil yang diurutkan dari kemiripan tertinggi
2. Pada fitur rekomendasi anime genre *iyashikei*. Jika tombol lihat analisis sentimen diklik akan berganti ke halaman menampilkan judul, jumlah *tweet*, dan persentase positif semua anime genre *iyashikei* dengan format tabel dan filterisasi berdasarkan jenis *library dictionary* atau perbandingan semua *library* terhadap 1 judul.
3. Setelah mengklik tombol *form* anime, halaman akan berpindah ke *form* dengan 6 pertanyaan yang wajib diisi. Setelah diisi, pengguna menekan tombol filter anime dan berpindah ke halaman baru yang menampilkan semua anime yang memenuhi isi *form*, diurutkan dari rating tertinggi dengan informasi judul, rating, genre, unsur, *link streaming*, dan *link* Anilist.

3.5.3 Mockup Website

Dengan 3 fitur utama, maka perlu rancangan UI agar pengguna dapat menemukan tiap fitur dengan mudah, penggunaan yang sederhana, dan nyaman untuk digunakan dalam waktu lama. Berikut *mockup* website:

Website Rekomendasi Anime

[←](#)
[→](#)
[↻](#)

Website Rekomendasi Anime
Temukan anime pilihanmu berdasarkan salah 1 fitur berikut
Anime dengan genre dan unsur termirip
Sentimen analisis khusus genre iyashikei
Form filterisasi anime

Cari Anime Termirip

Masukan Judul anime dan akan keluar rekomendasi anime dengan genre dan unsur cerita termirip

- Masukan judul Jepang atau Inggris
- Anime terelasi masih masuk dalam rekomendasi
- Contoh judul Jepang : Shingeki no Kyojin
- Contoh judul Inggris : Attack on Titan
- Contoh hasil : Shingeki no Kyojin 2, Shingeki no Kyojin: Chronicle

Contoh Anime Termirip

Anime	Termirip	Kemiripan(%)
Judul 1	Judul 2983	34.25
Judul 2	Judul 1425	28.11
Judul 3	Judul 92	20.85
Judul 4	Judul 8292	41.76
Judul 5	Judul 12433	30.81

Analisis Sentimen Anime Genre Iyashikei

Merupakan hasil perhitungan persentase positif anime genre iyashikei dengan data tweet anime terkait:

- Anime terpositif berdasarkan tweet
- Perbandingan 9 library analisis sentimen

Contoh Analisis Sentimen Anime (Vader)

Anime	Total	Positif
Yuru Camp△	1006	87.28
Non Non Biyori	969	83.38
Natsume Yuuichou	601	91.68
Gochuumon wa Usagi desu ka	145	94.48
Gakuen Babysitters	500	85.00

Filter Anime

Mencari anime yang spesifik? Gunakan filter untuk mencari yang cocok dengan kriteriamu:

Rating

Masukan angka dari 1-100

Jika memasukan 80, maka rating terendah yang tampil adalah 80

Fiksi

Ya : Fiksi / khayalan (Kekuatan super, hantu, dll)

Tidak : Non-fiksi / berbasis fakta

Apa saja: Tampilkan semua status

Status Karya Asli

Ya : Karya asli sudah tamat

Tidak : Karya asli masih berlanjut

Apa saja: Tampilkan semua status

Unsur LGBT+

Ya : Cerita memiliki unsur LGBT+

Tidak : Tidak ada unsur tersebut

Apa saja: Tampilkan semua status

Kekerasan Eskترم

Ya : Cerita erat dengan pembunuhan, pembantaian, genosida, dll

Tidak : Tidak ada unsur tersebut

Apa saja: Tampilkan semua status

Cerita Berat

Ya : Penonton perlu mengingat detail kejadian dan susah dipahami dalam sekali tonton

Tidak : Cerita ringan dan cocok untuk bersantai

Apa saja: Tampilkan semua status

Anime Recommendation System © 2025

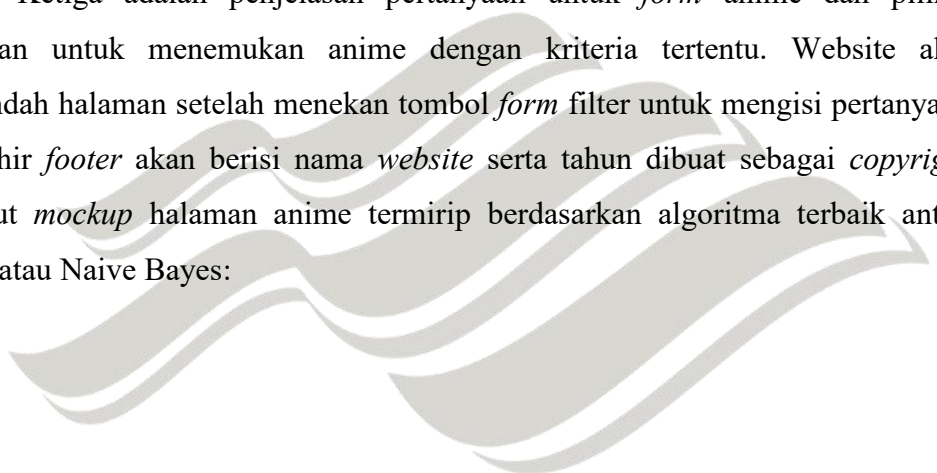
Gambar 3.5 Mockup landing page

39

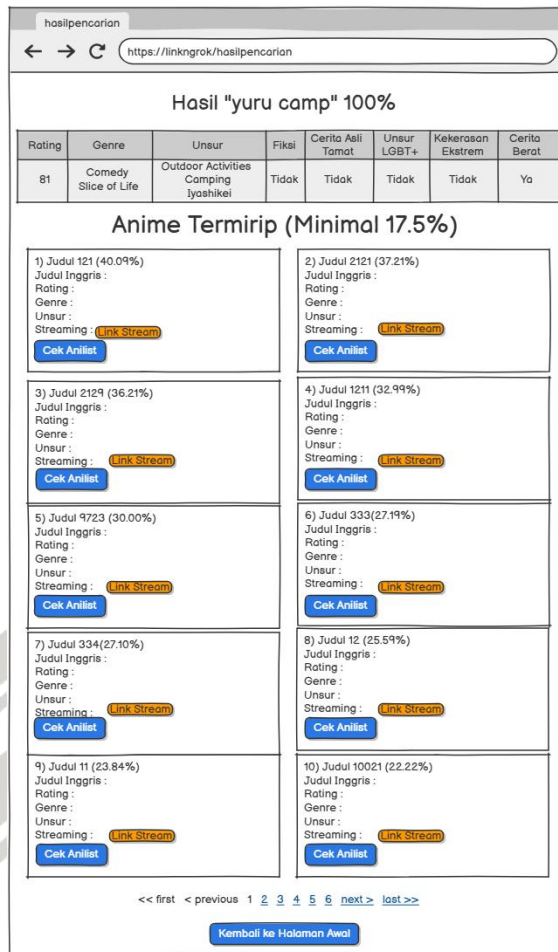
Pada gambar 3.5, *landing page* akan berisi *header* berisi nama *website* dan penjelasan 3 fitur yang disediakan. Fitur pertama adalah pencarian anime termirip menggunakan antara KNN atau Naive Bayes dengan confusion matrix terbaik dengan contoh hasil pada bagian kanan. Jika tombol cari ditekan setelah mengisi judul anime, *website* akan memberikan hingga 50 rekomendasi judul tercocok.

Kedua, analisis sentimen anime genre *iyashikei* dengan contoh hasil pada sebelah kanan. Setelah tombol lihat analisis sentimen ditekan, *website* akan berpindah ke halaman berisi tabel urutan anime dari persentase terpositif.

Ketiga adalah penjelasan pertanyaan untuk *form* anime dan pilihan jawaban untuk menemukan anime dengan kriteria tertentu. *Website* akan berpindah halaman setelah menekan tombol *form* filter untuk mengisi pertanyaan. Terakhir *footer* akan berisi nama *website* serta tahun dibuat sebagai *copyright*. Berikut *mockup* halaman anime termirip berdasarkan algoritma terbaik antara KNN atau Naive Bayes:

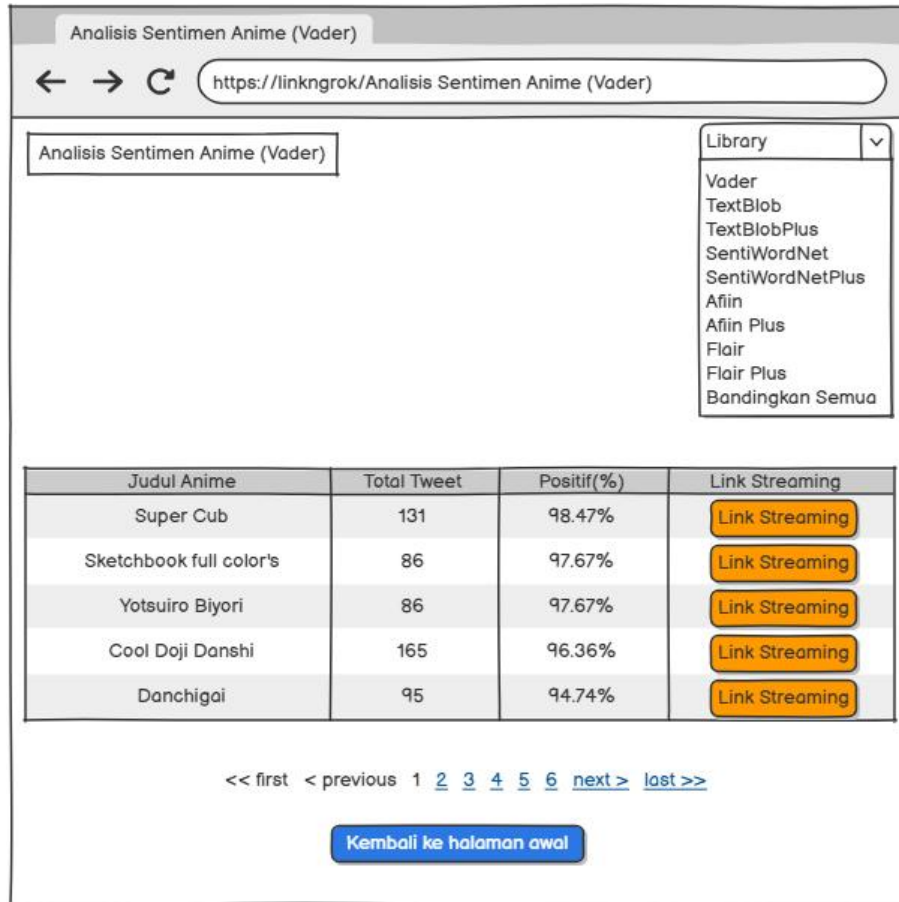


UNIVERSITAS
MA CHUNG



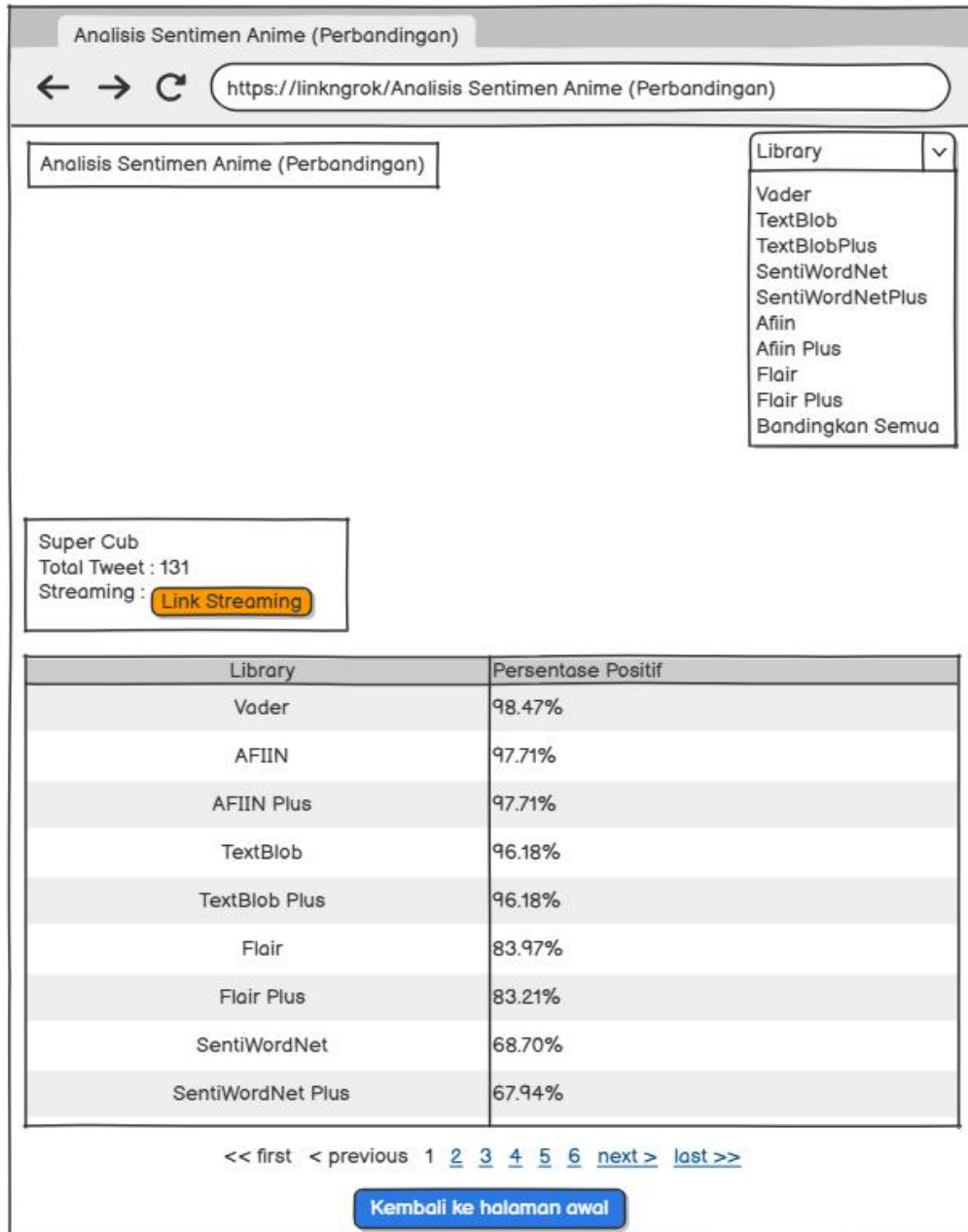
Gambar 3.6 *Mockup* Rekomendasi Anime Termirip Berdasarkan Algoritma Terbaik antara KNN atau Naive Bayes

Pada gambar 3.6, setelah pencocokan judul dengan database, akan dilakukan perhitungan berdasarkan algoritma terbaik antara KNN atau Naive Bayes untuk mencari hingga 50 anime dengan kemiripan tertinggi dengan batas minimal 17.5% kecocokan. Rekomendasi berisi judul, persentase kecocokan, judul Inggris, rating, genre, unsur, *link streaming*, dan *link Anilist*. Pada bagian bawah disediakan paginasi dan tombol kembali ke halaman awal. Berikut *mockup* halaman analisis sentimen:



Gambar 3.7 *Mockup* Analisis Sentimen Anime Genre *Iyashikei* Filter *library*

Pada gambar 3.7, halaman akan memiliki 10 filter di bagian kanan atas dengan pilihan *default* VADER. Selain bandingkan semua, 9 filter *library* akan memiliki tampilan sama 5 judul per halaman, total *tweet*, persentase positif, *link streaming*, paginasi, dan tombol kembali ke halaman awal. Untuk halaman analisis sentimen bandingkan semua dibuat dengan *mockup* sebagai berikut:



Gambar 3.8 *Mockup* Perbandingan *library* Analisis Sentimen Anime Genre *Iyashikei*

Pada gambar 3.8, halaman ini akan menampilkan 1 judul dengan perbandingan persentase positif dari 9 *library*, paginasi, dan tombol kembali ke halaman awal. Berikut *mockup* halaman *form* anime:

Filter Anime

← → ↻ <https://linkngrok/Filter Anime>

Filter Anime

Minimal Rating (1-100):

Misal 70

Apakah cerita fiksi? *

Pilih Opsi...

Ya
Tidak
Apa Saja

Cerita asli sudah tamat? *

Pilih Opsi...

Ya
Tidak
Apa Saja

Mengandung unsur LGBT+? *

Pilih Opsi...

Ya
Tidak
Apa Saja

Terdapat kekerasan ekstrem? *

Pilih Opsi...

Ya
Tidak
Apa Saja

Cerita berat? *

Pilih Opsi...

Ya
Tidak
Apa Saja

[Filter Anime](#) [Kembali ke Halaman Awal](#)

Gambar 3.9 *Mockup Form Filter Anime*

Pada gambar 3.9, halaman *form* akan berisi 6 pertanyaan yang wajib dijawab dengan pertanyaan rating diisi dengan angka dan 5 pertanyaan lagi memilih 1 opsi antara ya, tidak, dan apa saja yang telah dijelaskan pada halaman awal. Ketika sudah mengisi, klik filter anime untuk berpindah ke halaman filterisasi dan diberikan tombol untuk kembali ke halaman awal. Berikut *mockup* halaman hasil filterisasi anime:

3.6 Pengembangan Sistem

Setelah mendapatkan gambaran UI dan data yang dapat diambil, maka rancang bangun akan dikoding di Google Colab dengan flask, database bisa dihiraukan karena data excel yang ditaruh pada Google Colab bisa digunakan untuk filterasi sehingga tidak perlu dimasukan ke database seperti SQLite. Jika flask mengalami *error*, maka perlu menggunakan XAMPP dan database MySQL atau menggunakan ngrok untuk pembuatan portal *localhost*.

3.7 Pengambilan Data

Dalam pembuatan rancang bangun *website*, koding memiliki tahapan yaitu membuat excel sebagai database, pemodelan KNN Euclidian Distance, pembuatan tiap fitur, dan pembuatan *website* terutama UI yang dapat digunakan dengan mudah dan cepat oleh pengguna.

3.7.1 Pembuatan Excel sebagai database

Data yang akan diambil menggunakan API Anilist antara lain:

1. Judul anime dalam romaji dan Inggris

Digunakan untuk pengecekan ke database setelah pengguna memasukan judul pada *search box* untuk fitur pencarian anime termirip. Selain itu, judul romaji dan Inggris digunakan sebagai hasil rekomendasi anime termirip dan anime tercocok.

2. Link Anilist

Digunakan pada hasil anime termirip dan *form* anime jika pengguna masih membutuhkan informasi lebih lanjut karena keterbatasan data pada rekomendasi. Pengguna dapat menekan tombol menuju *website* Anilist.

3. Genre dan Unsur

Digunakan sebagai perhitungan kemiripan antar anime menggunakan hasil terbaik antara KNN dan Naive Bayes. Filter untuk menentukan fiksi atau non-fiksi, unsur LGBT+, unsur kekerasan, dan cerita berat yang dipakai untuk perhitungan kemiripan antar anime dan pertanyaan *form*. Kemudian ditampilkan semua genre dan 3 unsur dengan persentase tertinggi sebagai hasil pada pencarian anime termirip dan *form* anime.

4. Status karya asli

Dipakai untuk pertanyaan *form* anime, diambil dengan cara mengambil semua data berstatus *source*. Jika semuanya *finished* maka status karya asli tamat dan jika ada 1 yang *releasing* maka belum tamat.

5. Streaming Links

Link resmi untuk menonton anime seperti Cruchyroll, Youtube, dan lain-lain agar pengguna ketika sudah menemukan anime yang cocok bisa mengklik tombol menuju *website stream* anime terkait.

Setelah mengetahui data yang perlu diambil, maka koding dibuat dengan langkah-langkah berikut:

1. Mengimport dan *install library* yang diperlukan

Pengambilan data dilakukan melalui API Anilist menggunakan pustaka Requests untuk mengirimkan kueri GraphQL. Fungsi pengambilan data dirancang untuk menangani *rate limit* dengan mekanisme *retry* dan *exponential backoff*. Data disimpan dalam format Excel menggunakan Pandas, dengan progres pengambilan dipantau melalui Tqdm.

2. Pengumpulan data semua judul

Data diambil sesuai *query* yang diinginkan dan disesuaikan dengan perpindahan halaman sealur, misal anime *One Piece* dengan link “<https://anilist.co/anime/21/ONE-PIECE/>” terdapat *source* manga dengan link “<https://anilist.co/manga/30013/ONE-PIECE/>”, maka status manga dapat diambil. Jumlah id anime berjumlah 21.345 judul.

3. Membuat Link Anilist

Karena kurangnya data pada *website* rekomendasi, maka dapat menciptakan tombol menuju *website* Anilist yang spesifik per judul

```
1. def get_anilist_link(anime_id, title):
2.     """Buat AniList URL untuk tiap judul anime."""
3.     clean_title = ''.join(c if c.isalnum() or c in ('
    ', '-') else ' ' for c in title).strip()
```

```

4.     clean_title = clean_title.replace(' ', '-')
      ).replace('--', '-')
5. Return
      f"https://anilist.co/anime/{anime_id}/{clean_title}/"

```

Karena *link* https tidak ada pada API, maka perlu dibuat untuk semua judul dengan pola pada baris 5.

4. Pembuatan kriteria

Data yang tersedia pada Anilist tidak bisa langsung digunakan sebagai kriteria pada form, maka perlu fungsi untuk mengubah menjadi data *yes* dan *no*. Proses pembuatan kriteria menggunakan pengecekan ada tidaknya genre dan unsur tertentu. status sumber asli berdasarkan status *finished* atau *releasing* dari sumber seperti manga, *light novel*, atau novel. Fiksi / non-fiksi ditentukan berdasarkan genre seperti *Fantasy*, *Sci-fi*, dan lain-lain serta unsur seperti *magic*, *isekai*, dan lain-lain. Anime tanpa genre atau unsur tersebut dikategorikan sebagai non-fiksi. Unsur LGBT+: ditentukan berdasarkan dari genre seperti *Yuri* dan *Yaoi* serta unsur seperti *LGBTQ+ Themes*, *Shounen Ai*, dan lain-lain. Kekerasan keras ditentukan berdasarkan dari genre seperti *war*, *gore*, dan lain-lain serta unsur seperti *body horror*, *torture*, dan lain-lain. Cerita berat ditentukan berdasarkan dari genre seperti *psychological*, *mystery*, dan lain-lain serta unsur seperti *time manipulation*, *conspiracy*, dan lain-lain.

5. Pembuatan *Link Streaming*

Setiap anime memiliki tempat menonton yang berbeda sehingga perlu adanya *link* untuk menonton. Link yang dipakai hanya *website streaming* resmi seperti Crunchyroll, Netflix, Bilibili. Aniplus Asia, Ani-One Asia, iQIYI, Amazon Video, YouTube, dan lain-lain. Jika *link* menuju sosial media dan *website* yang bersifat pemberitahuan atau informasi seperti *official site*, Twitter, dan Instagram, maka tidak dipakai dalam *website*.

6. Simpan menjadi excel

Data yang dikumpulkan disimpan ke excel sebagai database dengan menambahkan pada setiap kolom nama fitur dengan format yang terstruktur dari kiri ke kanan adalah judul romaji, judul Inggris, *link* Anilist,

genre, unsur berserta persentasenya, rating, kriteria khusus seperti status karya asli, fiksi, unsur LGBT+, kekerasan ekstrem, apakah ceritanya berat, dan *link streaming*. File disimpan dengan nama *anime_data.xlsx* dan perlu diupload ke *session storage* sebagai data untuk website.

Data untuk *website* sudah didapatkan, namun untuk ditampilkan perlu dirapikan agar bisa digunakan seperti *streaming links* yang masih terhubung dan genre serta unsur yang perlu dihitung menggunakan KNN atau Naive Bayes tergantung hasil confusion matriks terbaik dari kedua algoritma. Jika terdapat data kosong seperti pada genre, unsur, dan rating perlu diperhatikan agar data kosong tidak merusak model.

3.7.2 Pembuatan Koding Anime Termirip Berdasarkan Genre dan Unsur

Sistem rekomendasi anime termirip dikembangkan menggunakan algoritma K-Nearest Neighbors (KNN) dengan jarak Euclidean dan Naive Bayes dengan cosine similarity untuk menghitung kemiripan berdasarkan genre, unsur, dan kriteria numerik. Pendekatan ini melibatkan perbandingan prediksi dan data aktual dengan membandingkan setiap data dalam database terhadap semua data lain. Proses pengembangan meliputi langkah-langkah berikut:

1. Persiapan *library*

Library seperti Pandas digunakan untuk memproses data tabel, NumPy untuk operasi numerik pada array multidimensi, `sklearn.neighbors.NearestNeighbors` untuk KNN, `sklearn.naive_bayes.MultinomialNB` untuk Naive Bayes, `MultiLabelBinarizer` dan `MinMaxScaler` untuk konversi teks ke biner dan normalisasi data (0-1), `cosine_similarity` untuk perhitungan jarak, `fuzzywuzzy` diinstal untuk menangani menanggulangi *typo* dengan *fuzzy matching*, `re` untuk normalisasi teks, `KFold` digunakan untuk *cross-validation* untuk mengecek overfitting dan underfitting, serta metrik seperti `accuracy_score`, `precision_recall_fscore_support`, dan `confusion_matrix` untuk evaluasi performa.

2. Pemrosesan data awal

Data dari excel `anime_data.xlsx` dimuat dan diproses untuk mengubah kolom teks yaitu status sumber asli, fiksi, LGBT+, kekerasan ekstrim, dan cerita berat menjadi angka biner (1/0) dan mengisi data rating kosong dengan *median*. Unsur diurai dari teks ke format numerik dengan bobot persentase, dan judul dinormalisasi dengan menghapus karakter non-alfanumerik.

3. Konversi Data ke Array Numerik

Data diubah ke matriks numerik dengan `MultiLabelBinarizer` untuk genre, pembuatan matriks bobot untuk unsur, dan normalisasi kolom numerik menggunakan `MinMaxScaler`. Fitur genre, unsur, dan numerik digabungkan untuk membentuk matriks fitur lengkap.

4. KNN

Algoritma KNN diterapkan dengan metrik Euclidean untuk menemukan 50 anime terdekat berdasarkan fitur. Pencocokan judul dilakukan dengan *fuzzy matching* dengan skor minimal 75% untuk menangani *typo*. Hasil disusun dalam `DataFrame` berisi judul, skor kemiripan, rating, genre, dan 3 unsur teratas.

5. Naive Bayes

Algoritma Naive Bayes digunakan dengan filter minimal 2 genre sama dan cosine similarity untuk rekomendasi. Rating dikelompokkan ke dalam kategori *Bad*, *Good*, *Very Good* untuk pelatihan model dengan probabilitas prediksi disesuaikan berdasarkan kemiripan fitur. Hasil berupa judul, skor kemiripan, rating, genre, dan 3 unsur teratas.

6. Evaluasi performa KNN dan Naive Bayes

Evaluasi menggunakan *10-fold cross-validation* dengan metrik akurasi, presisi, recall, F1-score, *average similarity* (rata-rata kemiripan judul dicari dengan rekomendasi 1 hingga rekomendasi ke n), *coverage* (persentase anime yang direkomendasikan dari semua judul di database), dan *diversity* (jarak antara rekomendasi 1 dengan 2, 1 sampai n, hingga n-1 sampai n). Data dibagi menjadi 90% latih dan 10% uji untuk mendeteksi overfitting atau underfitting. Nilai k pada KNN dites dari 1–9 dan kategori

rating 3 percobaan menjadi <60 jelek, 70–80 bagus, ≥ 80 sangat bagus, <65 jelek, 70–80 bagus, ≥ 80 sangat bagus <70 jelek, 70–80 bagus, ≥ 80 sangat bagus.

7. Fungsi *Main*

Fungsi *main* memuat data, memproses kueri judul misalnya *Shingeki no Kyojin*, dan menjalankan rekomendasi KNN dan Naive Bayes. Sistem menangani *typo* dengan *fuzzywuzzy* dan menyimpan hasil evaluasi, termasuk confusion matrix untuk analisis lebih lanjut.

3.7.3 Pembuatan Koding Analisis Sentimen

Analisis sentimen dikembangkan menggunakan Python di Google Colab untuk mengklasifikasikan *tweet* positif dan negatif terkait judul anime. Lima *library* yang digunakan yaitu VADER, TextBlob, SentiWordNet, AFINN, dan Flair. Data *tweet* mengandung *noise* seperti *emoticon*, *emoji*, *hashtag*, *link*, dan *tag username*, sehingga memerlukan pembersihan teks kecuali pada VADER yang memiliki fungsi bawaan dan diberi tambahan nama plus. Hasil dibagi menjadi sebelum dan sesudah pembersihan untuk mengevaluasi perbedaan akurasi. Analisis sentimen anime genre *iyashikei* dilakukan dengan langkah berikut:

1. Pengambilan Data

Tweet diambil menggunakan metode *tweet harvest* berbasis Node.js berdasarkan riset Halma Satria dengan autentikasi Twitter. Kueri spesifik dibuat dengan kombinasi kata kunci (AND/OR) dalam bahasa Inggris, menggunakan tanda kutip ganda (\\\\"judul\\") untuk menangani karakter khusus. Mode *latest* dipilih untuk relevansi *tweet* terkini dan diganti menjadi *top* jika data terlalu sedikit. Jumlah *tweet* yang bisa diambil 1000 *tweet* per 15 menit karena limitasi dari Twitter. File otomatis terbuat di dengan nama judul.csv dan *download* bersifat opsional.

2. Persiapan *library*

Library seperti Pandas untuk data tabel, NLTK untuk tokenisasi, NumPy untuk matriks, Seaborn dan Matplotlib untuk visualisasi, sklearn untuk TF IDF, StandardScaler, SVC, dan metrik evaluasi (akurasi, presisi, recall,

F1-score, confusion matrix). Library sentimen (VADER, TextBlob, SentiWordNet, AFINN, Flair) diinstal dan dikonfigurasi dengan pembersihan teks menggunakan *re* dan emoji.

3. Pemrosesan data

Tweet dari file *judul.csv* dibaca, data duplikat dihapus, dan teks dibersihkan seperti menghapus *tag*, *link*, *hashtag*, dan konversi *emoji*. Fitur TF-IDF dan CountVectorizer mengubah teks menjadi numerik, dilanjut StandardScaler dan SVC dengan GridSearchCV untuk optimasi parameter kernel, C, gamma.

4. Analisis sentimen

Setiap *library* menghitung skor sentimen jika ≥ 0 dianggap positif dan jika < 0 dianggap negatif. Dalam penentuan skor, VADER menggunakan SentimentIntensityAnalyzer, TextBlob dengan *polarity* namun tanpa *subjectivity* karena akan menyebabkan data berkurang, SentiWordNet dengan *pos_tag* manual, AFINN dengan skor langsung, dan Flair dengan tokenisasi eksternal karena tidak ada tokenisasi bawaan. Data dibagi 80% latih dan 20% uji menggunakan StratifiedKFold untuk 5-fold *cross-validation*.

5. Evaluasi performa

Menghitung akurasi, presisi, recall, F1-score, dan *confusion matrix*. Kesalahan klasifikasi yaitu false positive dan false negative dianalisis berdasarkan panjang *tweet* apakah terlalu pendek atau terlalu panjang. Visualisasi *wordcloud* dan *N-gram* 2-3 kata dibuat menggunakan TF-IDF untuk identifikasi faktor sentimen dan mengetahui kata positif dan negatif yang sering muncul pada *tweet* sehingga bisa merangkum pendapat pengguna terhadap judul anime seperti sifat karakter, anime terbaik, *ending* jelek, dan sebagainya.

6. Pengujian

Sistem diuji untuk mencari *bug*, *error* misalnya 0 *tweet* atau ketidakseimbangan kelas, dan performa. Statistik fold untuk mengecek ada

overfitting dan underfitting. Jika sudah lancar ulangi poin 1,3,4. dan 5 untuk setiap judul yang ingin dites karena menyatukan beberapa excel *tweet* dalam 1 excel tidak efisien.

3.8 Pembuatan Website Anime

Website dikembangkan menggunakan Flask untuk menyediakan fungsi perhitungan, akses database, dan UI sederhana. Proses meliputi:

1. Persiapan library

Instalasikan *fuzzywuzzy* untuk pencocokan teks, *Flask* untuk *framework website*, dan *pyngrok* untuk URL publik. Impor *Flask*, *request*, *render_template_string*, *redirect*, *url_for*, *jsonify*, *ngrok*, *numpy*, *pandas*, *re*, *NearestNeighbors*, *MultiLabelBinarizer*, dan *math.ceil* untuk mendukung operasi data dan UI.

2. Inisiasi awal

Inisiasi *Flask* dengan *static_folder='/content'* untuk akses gambar jika dibutuhkan dan konfigurasi *ngrok* untuk URL publik di *port 5000*.

3. Pemrosesan data

Muat *anime_data.xlsx* dan *sentimen_analisis.xlsx*, ubah teks ke numerik *Yes* ke 1 dan *No* ke 0, isi data kosong rating dengan *median* dan unsur dengan 0.5 jika tanpa persentase. Hapus baris kosong pada kolom kunci jika ditemukan. Normalisasi judul dan ekstrak *link streaming* dari kolom teks. Selain itu, *upload file* gambar jika dibutuhkan.

4. Fitur anime termirip berdasarkan genre dan unsur

Setelah memasukan *input* berupa judul oleh pengguna, dilakukan pencocokan dengan database *anime_data.xlsx* menggunakan *fuzzy matching* minimal 60% kecocokan, dikurangi dari konsep awal karena kemiripan 75% terlalu tinggi. Setelah judul yang dicari ditemukan di excel *anime_data.xlsx*, dilakukan perhitungan kemiripan menggunakan *K-Nearest Neighbors* dengan jarak euclidean untuk menghasilkan 10 rekomendasi berdasarkan genre, unsur, dan kriteria, dengan *one-hot encoding* untuk genre dan normalisasi fitur numerik. Kemudian tampilkan

rekomendasi berisi judul romaji, persentase kecocokan, judul Inggris, genre, 3 unsur teratas, *link streaming*, dan *link Anilist*.

5. Tabel Analisis Sentimen

Muat data sentimen dari *sentimen_analisis.xlsx*, isi tabel tergantung pilihan *library* yang disediakan pada kanan atas layar yaitu VADER, TextBlob, TextBlob plus, SentiWordNet, SentiWordNet plus, AFIIN, AFIIN plus, Flair, Flair plus, dan bandingkan semua (perbandingan 9 *library* dengan 1 judul). Tampilan hasil berisi judul, total tweet, persentase positif yang dihitung dari $(\text{True Positive} + \text{False Negative}) / \text{Tweet}$, dan *link streaming*.

6. Form Filter

Periksa isi 6 kriteria wajib yaitu rating, fiksi, status karya asli, LGBT+, kekerasan, cerita berat dari *input* pengguna dan filter anime sesuai data. Jika ada pertanyaan yang kosong keluarkan peringatan dan tidak bisa eksekusi filter. Jika terisi semua akan berpindah ke halaman asli dengan menampilkan pilihan pengguna di bagian atas layar. Setelah itu memunculkan 10 anime yang sesuai kriteria dengan urutan rating tertinggi berisi judul, genre, 3 unsur teratas, *link streaming*, dan *link Anilist*.

7. Desain UI

Seperti pada *mockup website*. Buat *landing page* dengan *header* bertuliskan judul dan 3 fitur, *body* berisi fitur pencarian anime termirip, contoh hasil pencarian anime termirip, fitur tabel analisis sentimen, contoh tabel analisis sentimen, penjelasan pertanyaan dan pilihan jawaban untuk form, serta *footer* berisi *trademark* dan kontak. Tambahkan *autocomplete* untuk saran pencarian sebanyak 5 judul dengan $\geq 60\%$ kecocokan *fuzzy matching*, paginasi untuk tabel sentimen, dan *alert* untuk form kosong.

8. Navigasi

Rute utama (/) mengarahkan ke /landing. Halaman hasil (pencarian, filter, sentimen) dirancang dengan detail anime, rekomendasi, atau tabel, lengkap dengan tombol kembali ke halaman awal.

9. Fungsi Main

Jalankan aplikasi dengan `app.run(host='0.0.0.0', port=5000)` untuk akses *localhost* dan ngrok untuk URL publik.

3.9 Evaluasi

Koding dan tampilan *website* perlu dievaluasi apakah telah berjalan sempurna tanpa *bug* dan mengeluarkan hasil yang sesuai

3.9.1 Evaluasi Metode Klasifikasi

1. SVM

Cek confusion matriks, accuracy, recall, precision, dan f-1 score apakah minimal mencapai 80%, Tiap fold dicek apakah mengalami overfitting atau underfitting.

2. KNN

Cek confusion matriks, accuracy, recall, precision, dan f-1 score apakah minimal mencapai 80%, Tiap fold dicek apakah mengalami overfitting atau underfitting

3.9.2 Evaluasi Website

1. *Black-box Testing*

Metode pengujian oleh pengembang dan bisa dibantu uji coba ke pengguna untuk memastikan bahwa aplikasi berfungsi sesuai dengan kebutuhan pengguna dan spesifikasi yang ditentukan. Meskipun tidak menguji koding secara langsung, metode ini efektif untuk mengidentifikasi masalah fungsional dan memvalidasi pengalaman pengguna. *Black-box testing* sering digunakan bersama dengan *white-box testing* untuk mencapai cakupan pengujian yang lebih komprehensif. Beberapa hal yang dilakukan pada *black-box testing* bisa secara manual atau otomatis menggunakan selenium, yaitu:

1. Menguji setiap fungsi aplikasi dengan memberikan *input* yang sesuai dan memverifikasi *output* terhadap persyaratan fungsional aplikasi.
2. Menguji setiap fungsi *entri* aplikasi untuk memeriksa semua titik.

3. Uji aliran layar GUI diperiksa sehingga pengguna dapat menavigasi seluruh aplikasi.
4. Menguji jika data diisi tidak memenuhi syarat atau tidak diisi jika bagian tersebut wajib, akan keluar *error*.
5. Tidak bisa memasuki halaman hasil tanpa melalui halaman sebelumnya.

2. Kecepatan

Dilihat jarak pergantian halaman terhitung dari klik submit, setiap *input* akan menghasilkan akan berbeda tergantung kekompleksan perbandingan data dan perhitungan.

3. Pendapat Pengguna

Pengguna dapat mengirimkan kritik dan saran terkait hal di atas, *bug*, kecepatan perpindahan halaman yang lama, tampilan UI, ukuran teks, isi yang susah dipahami, penambahan pertanyaan form, dan sebagainya.

3.10 Perbaikan

Apabila hasil kurang maksimal dan terdapat *bug*, maka koding perlu diperbaiki sebagai berikut:

3.10.1 White-box Testing

Cara pengujian perangkat lunak di mana penguji memiliki pengetahuan tentang kode secara manual atau otomatis sebagai berikut:

1. Memastikan setiap jalur melalui program dijalankan setidaknya satu kali.
2. Setiap titik keputusan dievaluasi sebagai benar atau salah setidaknya sekali.
3. Semua hasil potensial tercakup dengan mengeksekusi setiap titik keputusan dalam program setidaknya satu kali.
4. Pengujian iterasi banyak, nol, dan tunggal.
5. Menguji jalur penanganan kesalahan dan pengecualian program.
6. Pengujian kombinasi kondisi sederhana dan rumit di dalam titik-titik keputusan.
7. Mengevaluasi bagaimana perangkat lunak berperilaku ketika rentang input didorong hingga mencapai batasnya.

8. Menguji bagaimana berbagai modul atau komponen perangkat lunak berinteraksi satu sama lain.
9. Menentukan dan mengurangi kompleksitas siklomatik program.
10. Variabel yang digunakan tetapi tidak pernah didefinisikan.
11. Menemukan variabel yang didefinisikan tetapi tidak pernah digunakan.
12. Menemukan variabel yang didefinisikan beberapa kali sebelum digunakan.
13. Mengalokasikan variabel sebelum digunakan.

3.10.2 SVM

SVM merupakan algoritma yang membagi dataset ke dalam kelas menggunakan garis *hyperplane*, maka hal yang bisa diperbaiki adalah:

1. Jumlah data yang terlalu banyak pada 1 kelas tertentu.
2. Pastikan sudah normalisasi data.
3. Tidak ada nilai yang kosong dan duplikat.
4. *Hyperparamter tuning* pada nilai kernel, C, dan gamma.
5. Penambahan atau pengurangan fitur.
6. Cek *Confusion matriks*, *acuraccy*, *recall*, *precicion*, dan *f1-score*.
7. Buat K-fold *cross-validation* untuk mengecek overfitting dan underfitting.

3.10.3 KNN

KNN adalah metode klasifikasi dengan pengukuran jarak, maka beberapa hal yang perlu diperiksa apakah sudah berjalan dan menghasilkan nilai terbaik adalah berikut:

1. Pastikan sudah normalisasi data.
2. Tidak ada nilai yang kosong dan duplikat.
3. *Hyperparamter tuning* pada nilai k.
4. Penanganan nilai data tidak seimbang
5. Hilangkan fitur yang tidak relevan menggunakan PCA.
6. Cek *Confusion matriks*, *acuraccy*, *recall*, *precicion*, dan *f1-score*.
7. Buat K-fold *cross-validation* untuk mengecek overfitting dan underfitting.
8. Mengetes dengan metriks jarak lain seperti manhattan dan cosine

3.10.4 Naive Bayes

Naive Bayes adalah metode klasifikasi dengan pengukuran jarak, maka beberapa hal yang perlu diperiksa apakah sudah berjalan dan menghasilkan nilai terbaik adalah berikut:

1. Pastikan sudah normalisasi data.
2. Tidak ada nilai yang kosong dan duplikat.
3. Hyperparamter tuning varian Bayes yaitu Gaussian, Multinomial, atau Bernoulli
4. Penanganan nilai data tidak seimbang
5. Hilangkan fitur yang redundan
6. Cek Confusion matriks, acuraccy, recall, precicion, dan f1-score.
7. Buat K-fold *cross-validation* untuk mengecek overfitting dan underfitting.

3.10.5 Kecepatan

Keluarnya hasil tergantung dari seberapa cepat data diambil, dihitung, dan dibandingkan. Sehingga hal yang bisa diperbaiki adalah:

1. Pastikan koding berjalan efisien sesuai *white-box testing*.
2. Pengambilan dan filter data jika manual lewat koding bisa dimasukkan ke dalam database untuk mengurangi waktu.
3. Mencari cara perbandingan dengan data lain seminimal mungkin namun tetap benar.

3.10.6 Pendapat Pengguna

Tergantung pada kritik dan saran, kebanyakan sebagai berikut:

1. Pastikan UI rapi dan terbaca jelas.
2. Buat penjelasan jika ada istilah yang rumit.

BAB IV

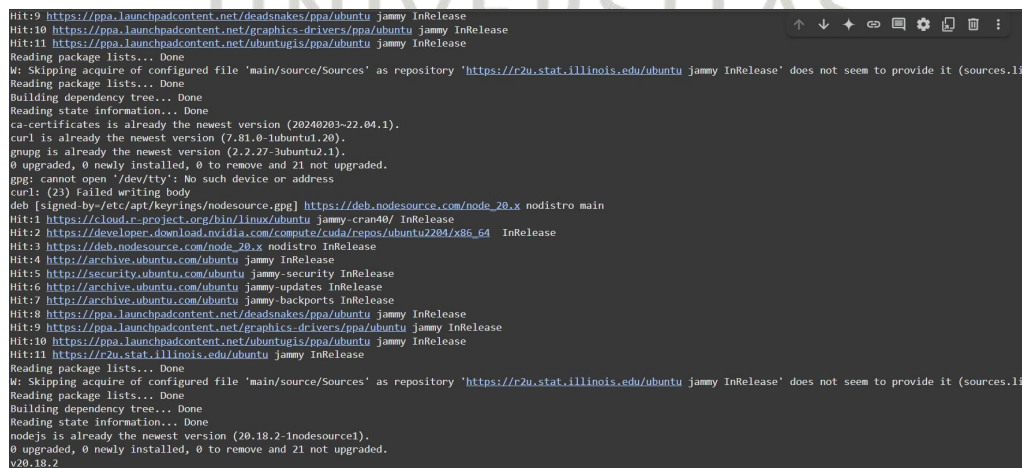
HASIL DAN PEMBAHASAN

4.1 Hasil Analisis Sentimen

Data *tweet* untuk analisis sentimen dikumpulkan dari Twitter/X menggunakan API dengan kata kunci terkait 79 judul anime *iyashikei*. Total *tweet* yang dianalisis bervariasi tiap judul dengan 18 judul tidak ada *tweet*. Jumlah *tweet* terendah yang bisa dianalisis adalah *Mokke* dan *Kokoro Toshokan* dengan 55 *tweet* dan tertinggi *Yuru Camp* dengan 1006 *tweet* dengan kebanyakan judul hanya memiliki 100–200 *tweet*. Data *tweet* dibersihkan dengan menghapus menggunakan koding standar *library* dan penghapusan tambahan seperti *username*, *link*, *hashtag*, dan mengubah *emoji* menjadi teks dengan label *plus*. Hasil analisis sentimen disimpan dalam file *sentimen_analisis.xlsx*. Hasil koding analisis sentimen sebagai berikut:

4.1.1 Install Node.js

Pada gambar 4.1 di bawah, indikasi *download* berhasil adalah keluarnya tulisan *node.js is already the newest version*. Pastikan koding dijalankan berjaga-jaga jika ada *update* terbaru pada *node.js*.

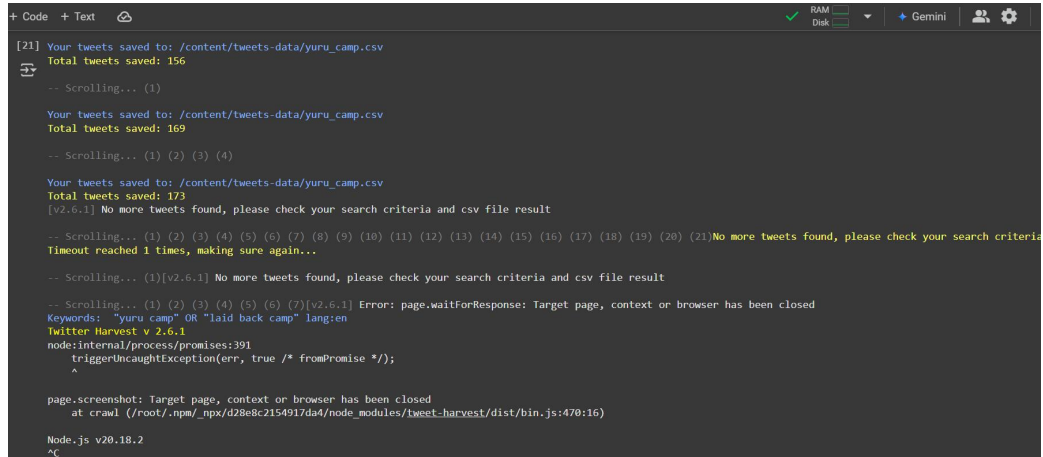


```
Hit:9 https://ppa.launchpadcontent.net/deadsnakes/ppa/ubuntu jammy InRelease
Hit:10 https://ppa.launchpadcontent.net/graphics-drivers/ppa/ubuntu jammy InRelease
Hit:11 https://ppa.launchpadcontent.net/ubuntugis/ppa/ubuntu jammy InRelease
Reading package lists... Done
W: Skipping acquire of configured file 'main/source/Sources' as repository 'https://r2u.stat.illinois.edu/ubuntu jammy InRelease' does not seem to provide it (sources.l
Reading package lists... Done
Building dependency tree... Done
Reading state information... Done
ca-certificates is already the newest version (20240203-22.04.1).
curl is already the newest version (7.81.0-1ubuntu1.20).
gnupg is already the newest version (2.2.27-3ubuntu2.1).
0 upgraded, 0 newly installed, 0 to remove and 21 not upgraded.
gpg: cannot open '/dev/pty': No such device or address
curl: (23) failed writing body
deb [signed-by=/etc/apt/keyrings/nodesource.gpg] https://deb.nodesource.com/node_20.x nodistro main
Hit:1 https://cloud.r-project.org/bin/linux/ubuntu jammy-cran40/ InRelease
Hit:2 https://developer.download.nvidia.com/compute/cuda/repos/ubuntu2204/x86_64 InRelease
Hit:3 https://deb.nodesource.com/node_20.x nodistro InRelease
Hit:4 http://archive.ubuntu.com/ubuntu jammy InRelease
Hit:5 http://security.ubuntu.com/ubuntu jammy-security InRelease
Hit:6 http://archive.ubuntu.com/ubuntu jammy-updates InRelease
Hit:7 http://archive.ubuntu.com/ubuntu jammy-backports InRelease
Hit:8 https://ppa.launchpadcontent.net/deadsnakes/ppa/ubuntu jammy InRelease
Hit:9 https://ppa.launchpadcontent.net/graphics-drivers/ppa/ubuntu jammy InRelease
Hit:10 https://ppa.launchpadcontent.net/ubuntugis/ppa/ubuntu jammy InRelease
Hit:11 https://r2u.stat.illinois.edu/ubuntu jammy InRelease
Reading package lists... Done
W: Skipping acquire of configured file 'main/source/Sources' as repository 'https://r2u.stat.illinois.edu/ubuntu jammy InRelease' does not seem to provide it (sources.l
Reading package lists... Done
Building dependency tree... Done
Reading state information... Done
nodejs is already the newest version (20.18.2-nodeSource1).
0 upgraded, 0 newly installed, 0 to remove and 21 not upgraded.
v20.18.2
```

Gambar 4.1 Instalasi Node.js berhasil

4.1.2 Tweet Harvest

Saat menjalankan *harvest*, *tweet* akan tersimpan maksimal 1 kali per *scrool* dan ada kemungkinan data di bawah *limit* yang ditetapkan.



```
[21] Your tweets saved to: /content/tweets-data/yuru_camp.csv
Total tweets saved: 156

-- Scrolling... (1)

Your tweets saved to: /content/tweets-data/yuru_camp.csv
Total tweets saved: 169

-- Scrolling... (1) (2) (3) (4)

Your tweets saved to: /content/tweets-data/yuru_camp.csv
Total tweets saved: 173
[v2.6.1] No more tweets found, please check your search criteria and csv file result

-- Scrolling... (1) (2) (3) (4) (5) (6) (7) (8) (9) (10) (11) (12) (13) (14) (15) (16) (17) (18) (19) (20) (21) No more tweets found, please check your search criteria
Timeout reached 1 times, making sure again...

-- Scrolling... (1)[v2.6.1] No more tweets found, please check your search criteria and csv file result

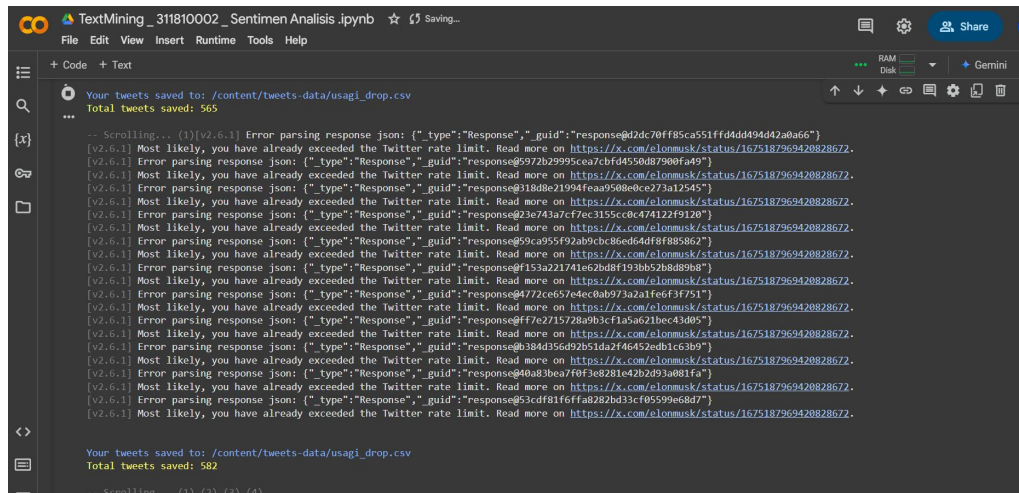
-- Scrolling... (1) (2) (3) (4) (5) (6) (7)[v2.6.1] Error: page.waitForResponse: Target page, context or browser has been closed
Keywords: "yuru camp" OR "laid back camp" lang:en
Twitter Harvest v 2.6.1
node:internal/process/promises:391
    triggerUncaughtException(err, true /* fromPromise */);
    ^

page.screenshot: Target page, context or browser has been closed
at crawl (/root/.npm/_npx/d28e8c2154917da4/node_modules/tweet-harvest/dist/bin.js:470:16)

Node.js v20.18.2
```

Gambar 4.2 *Harvest No More Tweet Found dan Time Out*

Pada gambar 4.2, muncul tulisan *no more tweets found* dan *time out reached* untuk melakukan *scrolling* ulang dan berhenti jika tidak ditemukan *tweet* lagi. Hal ini menunjukkan jika judul yang dicari kurang populer.



```
TextMining_311810002_Sentimen Analisis.ipynb
File Edit View Insert Runtime Tools Help

+ Code + Text

Your tweets saved to: /content/tweets-data/usagi_drop.csv
Total tweets saved: 565

-- Scrolling... (1)[v2.6.1] Error parsing response json: {"_type":"Response","_guid":"response@d2dc70ff85ca551ffdd44d42a0a66"}
[v2.6.1] Most likely, you have already exceeded the Twitter rate limit. Read more on https://x.com/elonmusk/status/1675187969420828672.
[v2.6.1] Error parsing response json: {"_type":"Response","_guid":"response@5972b29995cea7cbfd4550d87900fa49"}
[v2.6.1] Most likely, you have already exceeded the Twitter rate limit. Read more on https://x.com/elonmusk/status/1675187969420828672.
[v2.6.1] Error parsing response json: {"_type":"Response","_guid":"response@318b0e21994fead9080dec273a13545"}
[v2.6.1] Most likely, you have already exceeded the Twitter rate limit. Read more on https://x.com/elonmusk/status/1675187969420828672.
[v2.6.1] Error parsing response json: {"_type":"Response","_guid":"response@23e743a7c7fec3155cc0474122f9120"}
[v2.6.1] Most likely, you have already exceeded the Twitter rate limit. Read more on https://x.com/elonmusk/status/1675187969420828672.
[v2.6.1] Error parsing response json: {"_type":"Response","_guid":"response@59ca955f92ab9c86ed64d8f885862"}
[v2.6.1] Most likely, you have already exceeded the Twitter rate limit. Read more on https://x.com/elonmusk/status/1675187969420828672.
[v2.6.1] Error parsing response json: {"_type":"Response","_guid":"response@f53a221741e62b08f193b052b080b0b"}
[v2.6.1] Most likely, you have already exceeded the Twitter rate limit. Read more on https://x.com/elonmusk/status/1675187969420828672.
[v2.6.1] Error parsing response json: {"_type":"Response","_guid":"response@4772ce657e4ec0ab973a2a1fe6f3f751"}
[v2.6.1] Most likely, you have already exceeded the Twitter rate limit. Read more on https://x.com/elonmusk/status/1675187969420828672.
[v2.6.1] Error parsing response json: {"_type":"Response","_guid":"response@ff7e2715728a9b3cfa5a621bec43d05"}
[v2.6.1] Most likely, you have already exceeded the Twitter rate limit. Read more on https://x.com/elonmusk/status/1675187969420828672.
[v2.6.1] Error parsing response json: {"_type":"Response","_guid":"response@b084356d92b51dc2f46a57ed01c6b09"}
[v2.6.1] Most likely, you have already exceeded the Twitter rate limit. Read more on https://x.com/elonmusk/status/1675187969420828672.
[v2.6.1] Error parsing response json: {"_type":"Response","_guid":"response@40a83bea7f0f3e8281e42b2d93a081fa"}
[v2.6.1] Most likely, you have already exceeded the Twitter rate limit. Read more on https://x.com/elonmusk/status/1675187969420828672.
[v2.6.1] Error parsing response json: {"_type":"Response","_guid":"response@53cdf81f6ffa8282bd33cf05599e68d7"}
[v2.6.1] Most likely, you have already exceeded the Twitter rate limit. Read more on https://x.com/elonmusk/status/1675187969420828672.

Your tweets saved to: /content/tweets-data/usagi_drop.csv
Total tweets saved: 582

-- Scrolling... (1) (2) (3) (4)
```

Gambar 4.3 *Harvest Melebihi Limit tapi Berjalan Terus*

Pada gambar 4.3, *harvest* mencapai batas meskipun di bawah 1000, akan muncul tulisan *error parsing* dan *already exceeded the Twitter rate limit*. Jika

masih ditemukan *tweet*, *harvest* akan dilanjutkan sampai *limit* yang ditentukan pada koding.

```
-- Scrolling... (1)
Your tweets saved to: /content/tweets-data/usagi_drop.csv
Total tweets saved: 958

-- Taking a break, waiting for 10 seconds...

-- Scrolling... (1) (2) (3)

Your tweets saved to: /content/tweets-data/usagi_drop.csv
Total tweets saved: 972

-- Scrolling... (1)

Your tweets saved to: /content/tweets-data/usagi_drop.csv
Total tweets saved: 985

-- Scrolling... (1) (2) (3) (4)

Your tweets saved to: /content/tweets-data/usagi_drop.csv
Total tweets saved: 999

-- Scrolling... (1)

Your tweets saved to: /content/tweets-data/usagi_drop.csv
Total tweets saved: 1016
Got 1016 tweets, done scrolling...
```

Gambar 4.4 *Harvest Melebihi Limit*

Gambar 4.4, ketika *tweet* mencapai *limit* masih bisa bertambah sekali yaitu ketika pada angka terakhir sebelum *limit* akan bertambah maksimal 20. Terakhir, akan muncul jumlah *tweets*, *done scrolling*.

4.1.3 Fold

Nilai fold wajib diperhatikan agar tidak terjadi overfitting dan underfitting. Perubahan terutama pada akurasi jika terlalu jauh seperti 10%, maka perlu dilakukan *hyperparameter-tuning* atau terdapat kesalahan pada data yang kurang bagus atau kurangnya pembersihan.

```
[nlTK_data] Downloading package vader_lexicon to /root/nltk_data...
[nltk_data] Package vader_lexicon is already up-to-date!
[nltk_data] Downloading package opinion_lexicon to /root/nltk_data...
[nltk_data] Package opinion_lexicon is already up-to-date!
Best C value: 0.1
Best kernel: poly
Best gamma: scale
Fold 1 Metrics:
Accuracy: 0.86, Precision: 0.86, Recall: 1.00, F1-Score: 0.93
Sentiment Cases: {'fold': 1, 'actual_positive_predicted_positive': 138, 'actual_positive_predicted_negative': 0, 'actual_negative_predicted_positive': 22, 'actual_n
Fold 2 Metrics:
Accuracy: 0.87, Precision: 0.87, Recall: 1.00, F1-Score: 0.93
Sentiment Cases: {'fold': 2, 'actual_positive_predicted_positive': 138, 'actual_positive_predicted_negative': 0, 'actual_negative_predicted_positive': 21, 'actual_n
Fold 3 Metrics:
Accuracy: 0.86, Precision: 0.86, Recall: 1.00, F1-Score: 0.93
Sentiment Cases: {'fold': 3, 'actual_positive_predicted_positive': 138, 'actual_positive_predicted_negative': 0, 'actual_negative_predicted_positive': 22, 'actual_n
Fold 4 Metrics:
Accuracy: 0.87, Precision: 0.87, Recall: 1.00, F1-Score: 0.93
Sentiment Cases: {'fold': 4, 'actual_positive_predicted_positive': 138, 'actual_positive_predicted_negative': 0, 'actual_negative_predicted_positive': 21, 'actual_n
Fold 5 Metrics:
Accuracy: 0.87, Precision: 0.87, Recall: 1.00, F1-Score: 0.93
Sentiment Cases: {'fold': 5, 'actual_positive_predicted_positive': 138, 'actual_positive_predicted_negative': 0, 'actual_negative_predicted_positive': 20, 'actual_n
```

Gambar 4.5 Nilai C, Kernel, Gamma, Fold Terbaik, dan Hasil Tiap *fold*

Gambar 4.5 menunjukan hasil yang bagus karena jarak antar statistik berdekatan dan tinggi tiap fold dengan $C=0.1$, $\text{kernel}=\text{poly}$, dan $\text{gamma}=\text{scale}$,

akurasi terendah 86% dan tertinggi 87%. Precision dan recall yang tinggi menunjukkan SVC cenderung menganggap data apapun sebagai positif.

4.1.4 Hasil Analisis Sentimen Tweet

Setelah *cross-validation*, akan dimunculkan *tweet*, prediksi dari *library*, dan aktual dari SVC. Hasil analisis sentimen tweet sebagai berikut:

1. VADER

True Positive Example:

Text: @claire_yuri_alt Yeah it's not like usagi drop or smth lol you can kinda smell the incest at a glance. It's almost like the ppl who were surprised coffin of andy and leyley had incest there's a certain vibe you learn to pick up on that isn't always canonized but often is.

Actual Label: Positive

Predicted Label: Positive

2. TextBlob

True Positive Example:

*Text: OH GOD NO NO!!!
NOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOO! ANOTHER USAGI
DROP NOOOOOOOOOOOOOOOOOOOOOOOOOOOOO!!!! WHY!? THAT
STARTED SO GREAT MAN WHY!? <https://t.co/4BANJvNNuw>*

Actual Label: Positive

Predicted Label: Positive

3. TextBlob + Pembersihan Data

True Positive Example:

Text: Usagi drop bro keep dropping his preferred age dawgggg

Actual Label: Positive

Predicted Label: Positive

4. SentiWordNet

True Positive Example:

*Text: @Kane0Lore Hyouka Kino's Journey (2003) Land of the Lustrous
Moribito Usagi Drop (Thank God) Shirobako Space Brothers*

Actual Label: Positive

Predicted Label: Positive

5. SentiWordNet + Pembersihan Data

True Positive Example:

Text: apparently there are a lot of japanese media where a guy adopts a girl and they end up getting married. usagi drop was another one. seriously disgusting.

Actual Label: Positive

Predicted Label: Positive

6. AFIIN

True Positive Example:

Text: @Nijigakupriv It came out while I was busy af lol. Still catching up with that banished hero anime Close enough welcome back Usagi drop

Actual Label: Positive

Predicted Label: Positive

7. AFIIN + Pembersihan Data

True Positive Example:

Text: ok so true. actors songs connection is a music anime its peak. theres adachi to shimamura...yuri. aharen san wa hakaranai is a comedy. amagi brilliant park is a comedy. kakushigoto is a comedy. sakamoto desu ga is a comedy. usagi drop is a cute family story

Actual Label: Positive

Predicted Label: Positive

8. Flair

True Positive Example:

Text: GM/GN. Usagi Drop is now on @KreatePlatform (https://t.co/BqMt9gkT3P) - 1/1 E 50 A - thank you and have a nice weekend! https://t.co/6Qny6YIN4J

Actual Label: Positive

Predicted Label: Positive

9. Flair + Pembersihan Data

True Positive Example:

Text: Usagi drop started off wholesome and look how that ended. So don t attack ppl for being apprehensive with a weird scenario and title like this show has it s understandable and there quite a few other manga I could mention as well. I glad to know this one doesn t go that way.

Actual Label: Positive

4.1.5 Confusion Matrix

Confusion matrix berisi accuracy, precision, recall, dan f1-score berdasarkan hasil true positive, false negative, false positive, dan true negative. Berikut contoh perbandingan hasil untuk pencarian *Usagi Drop*:



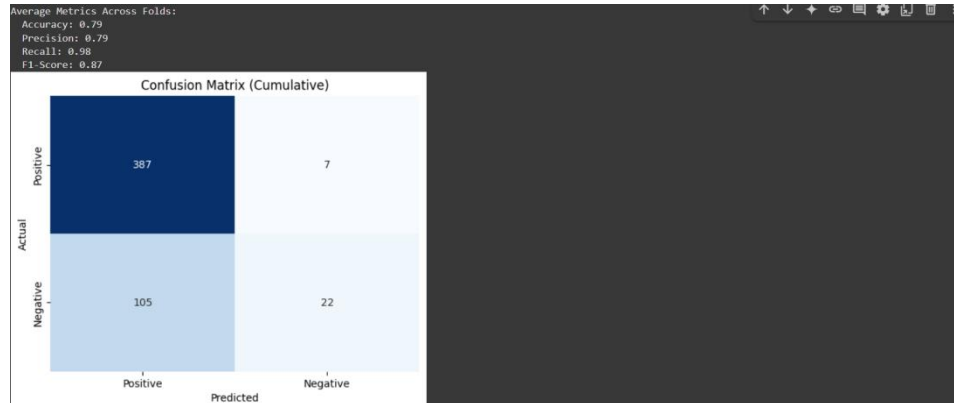
Gambar 4.6 Confusion Matrix VADER *Usagi Drop*

Pada gambar 4.6, VADER mendeteksi lebih banyak aktual negatif dibanding positif. Secara akurasi cukup 69%.



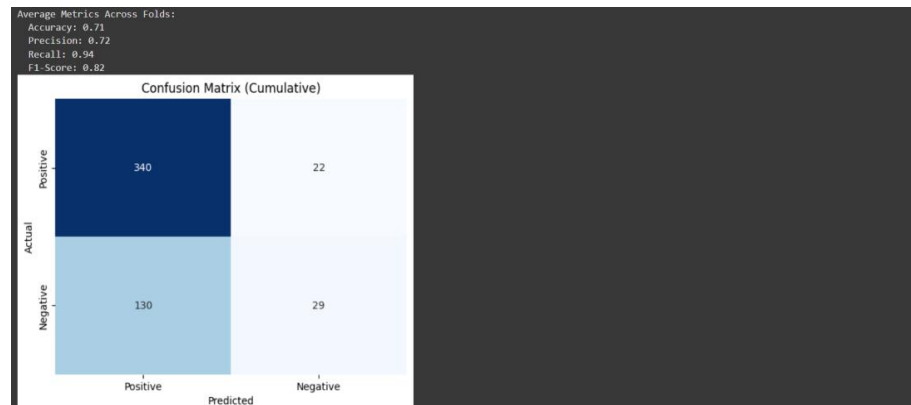
Gambar 4.7 Confusion Matrix TextBlob *Usagi Drop*

Pada gambar 4.7, TextBlob memiliki accuracy 81% dan recall 98% menunjukkan data cenderung ke positif. Hal ini terjadi karena TextBlob dijalankan tanpa filter *subjectivity* sehingga text dengan *subjectivity* 0 dan *polarity* 0 akan dianggap positif dalam aturan positif jika *polarity* ≥ 0 .



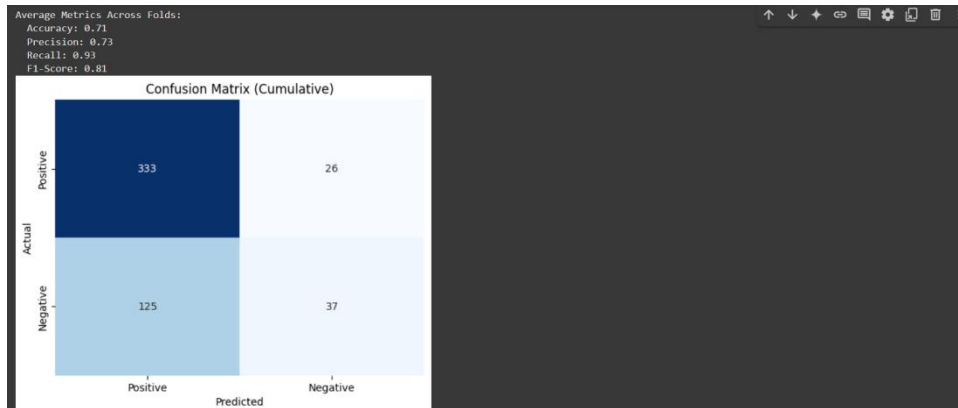
Gambar 4.8 Confusion Matrix TextBlob + Pembersihan Data *Usagi Drop*

Pada gambar 4.8, TextBlob dengan pembersihan data memiliki accuracy 80% dan recall 98% menunjukkan data cenderung ke positif. Akurasi menurun sedikit karena pembersihan teks membuat text semakin banyak positif.



Gambar 4.9 Confusion Matrix SentiWordNet *Usagi Drop*

Pada gambar 4.9, SentiWordNet cenderung ke arah positif dengan recall 94% dan akurasi 77%. SentiWordNet lebih banyak aktual data negatif dibanding TextBlob tetapi cenderung memprediksi sebagai positif membuat akurasinya lebih rendah.



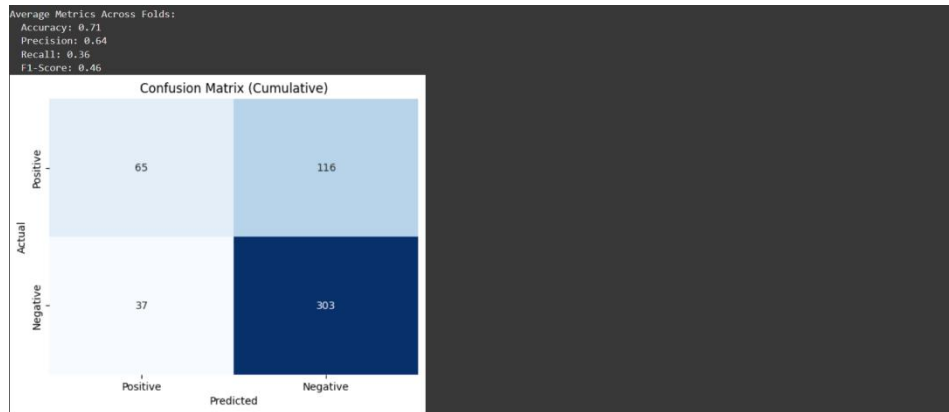
Gambar 4.10 Confusion Matrix SentiWordNet + Pembersihan Data *Usagi Drop*

Pada gambar 4.10, SentiWordNet dengan pembersihan data memiliki accuracy 76% dan recall 93% menunjukkan data cenderung ke positif. Akurasi menurun sedikit karena data positif mulai diprediksi sebagai negatif.



Gambar 4.11 Confusion Matrix AFINN *Usagi Drop*

Pada gambar 4.11, AFINN memiliki data aktual mayoritas negatif dengan akurasi sebesar 75%, lebih bagus dari VADER jika data positif dan negatif tercampur.



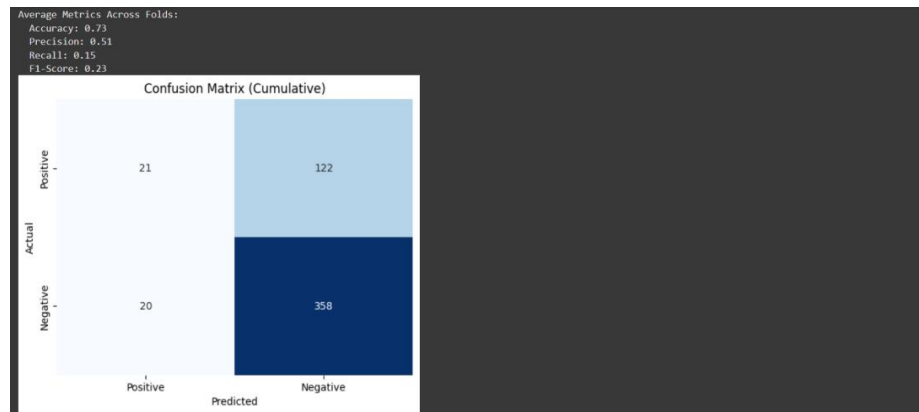
Gambar 4.12 Confusion Matrix AFINN + Pembersihan Data *Usagi Drop*

Pada Gambar 4.12, pembersihan data pada AFINN menyebabkan prediksi menurun, justru semakin banyak data yang salah klasifikasi.



Gambar 4.13 Confusion Matrix Flair Usagi Drop

Pada gambar 4.13, mayoritas data aktual Flair negatif dan prediksi cenderung ke negatif. Secara akurasi masih bagus 72% namun recall dan f1-score jelek di angka 55% dan 23%.



Gambar 4.14 Confusion Matrix Flair + Pembersihan Data *Usagi Drop*

Pada gambar 4.14. Flair mulai bisa memprediksi data positif sehingga peningkatan pada *recall* dan *f1-score* walaupun secara akurasi tetap 72%.

4.1.6 Error

Kesalahan berupa false positive dan false negative wajib diperhatikan agar bernilai kecil. Beberapa faktor lain bisa memengaruhi antara lain *tweet* yang terlalu pendek dan panjang.



Gambar 4.15 Error SentiWordNet

Pada gambar 4.15, terdapat 6 *tweet* dengan jumlah kata ≤ 6 , 20 *tweet* ≥ 280 , false positive 166, dan false negative 43. Kesalahan seperti ini wajib diperbaiki terutama jika data rusak dengan cara pembersihan data dan *hyperparameter-tuning*.



Gambar 4.16 Error SentiWordNet + Pembersihan Data

Pada gambar 4.16, jumlah *tweet* ≤ 6 kata tersisa 14 data dan *tweet* ≥ 280 karakter hanya 1, menunjukkan pembersihan data berhasil. False negative naik menjadi 47 dan False positive meningkat menjadi 180. Jumlah *error* meningkat hanya berbeda jenis. Jumlah teks terlalu pendek bisa saja meningkat karena pada *tweet* terlalu panjang mengalami banyak penghapusan. Sebaliknya *tweet* terlalu pendek jika memiliki *emoticon* yang dijadikan kata akibat fitur deemoji menyebabkan jumlah kata bertambah.

4.1.7 Wordcloud 1 Kata

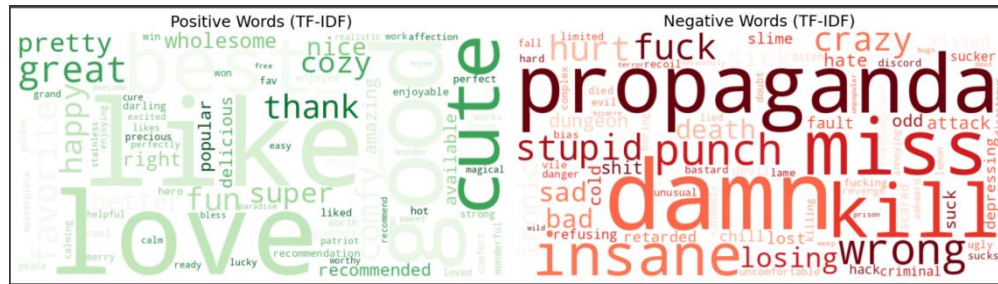
Kata positif dan negatif yang paling memengaruhi memiliki skor *Term Frequency-Inverse Document Frequency* (TF-IDF) tinggi. TF-IDF mengukur seberapa penting sebuah kata pada sebuah dokumen.

```
Top 10 Most Used Positive Words based on TF-IDF:
like: 15.6632
love: 14.1688
good: 18.2972
best: 8.4781
cute: 7.7049
great: 6.3178
favorite: 5.5017
thank: 4.9446
comfy: 4.7342
fun: 4.5733

Top 10 Most Used Negative Words based on TF-IDF:
damn: 1.3972
propaganda: 1.0870
miss: 0.9906
kill: 0.9870
insane: 0.9739
punch: 0.9154
wrong: 0.9030
stupid: 0.8445
fuck: 0.8339
hurt: 0.7915
```

Gambar 4.17 10 Teratas Kata Positif dan Negatif

Pada gambar 4.17. kata yang menunjukkan kesukaan akan judul anime adalah *like, love, good, best, great, dan favorites*, faktor cerita seperti *fun, cute dan comfy*. *Thank* kekurangan konteks jika hanya melihat 1 kata. Untuk kata negatif hanya *stupid dan fuck* yang menjelaskan. *Hurt, wrong, punch, kill, miss, dan propaganda* kurang jelas hubungannya dengan anime. Sementara itu *insane* dan *miss* bisa bermakna positif atau negatif.



Gambar 4.18 Wordcloud 1 Kata Positif dan Negatif

Wordcloud wajib dipisah agar dapat mencari tau dengan cepat faktor yang menyebabkan positif dan negatif. Terlihat pada gambar 4.18 kata negatif sangat banyak tetapi secara skor tidak sebesar positif. Pemakaian kata negatif lebih variatif tetapi jumlah penggunaan tidak seintens positif.

4.1.8 Wordcloud 2-3 kata (N-gram)

Wordcloud 1 kata apabila kurang jelas faktor atau konteksnya dapat menggunakan *N-gram*, N kata berurut yang bisa dipakai untuk mencari ringkasan baik positif dan negatif. *N-Gram* teroptimal adalah 2 hingga 3 karena di atas 3 kata sudah jarang pemakaiannya sehingga skor TF-IDF akan rendah.



Gambar 4.19 N-Gram Teratas Kata Positif dan Negatif

Pada gambar 4.19, *shiny days* memang positif secara *library* namun dalam konteks anime kata tersebut adalah judul lagu. Kata negatif masih banyak kesalahan, *the fuck* bisa saja ekspresi terkejut yang positif, *cold weather* tidak berhubungan langsung dengan anime, 8 kata negatif sisanya merupakan judul anime. Gabungan kata terkadang membuat subjek dan objek yang netral dianggap sebagai kata positif atau negatif.



Gambar 4.20 Worcloud 2-3 Kata Positif dan Negatif

Gambar 4.20 menunjukkan visual *N-Gram* dalam *wordcloud*, akan tetapi hanya ditampilkan 10 teratas saja.

4.1.9 Perbandingan Akurasi 9 Sentimen Analisis

Penelitian ini menggunakan 5 *library*, VADER, TextBlob, SentiWordNet, AFINN, dan Flair. 4 *library* selain VADER ditambahkan pembersihan teks agar terlihat perubahan statistik sebelum dan sesudah adanya pembersihan. Berikut adalah hasil dari 9 *library* analisis sentiment tiap *keyword*:

Tabel 4.1 Perbandingan hasil dari 10 judul anime genre *iyashikei* terpopuler

Keyword	Library	Accur acy	Preci sion	Recall	F1	TP	FN	FP	TN
Mushishi Latest	Vader	83	83	98	80	348	7	71	38
	TextBl ob	89	89	100	94	411	0	53	0
	TextBl ob Plus	88	89	100	94	410	1	53	0
	SentiW ordNet	71	73	96	83	318	14	120	12
	SentiW ordNet Plus	73	74	96	84	322	12	115	15

Keyword	Library	Accuracy	Precision	Recall	F1	TP	FN	FP	TN
	AFIIN	85	85	100	92	392	0	70	2
	AFIIN Plus	85	85	100	92	391	1	70	2
	Flair	78	78	98	87	339	6	98	21
	Flair Plus	77	78	97	86	338	12	94	20
Barakam on Latest	Vader	84	84	98	90	366	7	72	40
	TextBlob	89	89	100	94	431	0	54	0
	TextBlob Plus	89	89	100	94	431	0	54	0
	SentiWordNet	80	80	99	89	375	4	92	14
	SentiWordNet Plus	80	80	99	89	377	2	93	13
	AFIIN	85	85	100	92	403	0	71	11
	AFIIN Plus	85	85	100	92	403	0	71	11
	Flair	83	85	93	89	343	24	60	58
	Flair Plus	83	85	93	89	344	26	61	54
Natsume Yuujinchou Natsume's Book of Friends Latest	Vader	93	93	100	96	551	0	42	8
	TextBlob	91	91	100	95	539	0	54	8
	TextBlob Plus	91	91	100	95	539	0	54	8
	SentiWordNet	85	85	100	92	504	2	87	8
	SentiWordNet Plus	85	85	99	92	502	4	86	9
	AFIIN	91	91	100	95	540	0	52	9
	AFIIN Plus	91	91	100	95	538	2	53	8
	Flair	82	82	100	90	491	0	110	0
	Flair Plus	81	82	99	89	484	7	109	1
Yuru Camp Laid Back Camp Latest	Vader	88	88	100	94	887	1	115	13
	TextBlob	90	90	100	95	896	2	102	6
	TextBlob Plus	89	89	100	94	898	0	108	0
	SentiWordNet	74	77	92	84	664	56	202	84

Keyword	Library	Accuracy	Precision	Recall	F1	TP	FN	FP	TN
	SentiWordNet Plus	77	77	99	87	744	10	220	32
	AFIIN	90	91	100	95	900	2	94	10
	AFIIN Plus	90	90	100	95	898	4	95	9
	Flair	76	76	100	86	759	2	240	5
	Flair Plus	77	77	100	87	765	1	234	6
Usagi Drop Latest	Vader	69	56	59	57	110	75	86	250
	TextBlob	78	79	98	87	385	9	105	22
	TextBlob Plus	79	79	98	87	387	7	105	22
	SentiWordNet	71	72	94	82	340	22	130	29
	SentiWordNet Plus	71	73	93	81	333	26	125	37
	AFIIN	71	59	54	56	97	84	66	274
	AFIIN Plus	71	64	36	46	65	116	37	303
	Flair	72	51	20	28	19	119	27	346
	Flair Plus	73	51	15	23	21	122	20	358
Tanaka-kun wa Itsumo Kedaruge Tanaka-kun is Always Listless Latest	Vader	87	87	100	93	480	1	71	87
	TextBlob	88	89	100	94	489	1	63	88
	TextBlob Plus	88	89	100	94	489	0	65	88
	SentiWordNet	85	85	100	92	469	2	82	85
	SentiWordNet Plus	85	85	100	92	469	2	82	85
	AFIIN	89	89	100	94	491	0	63	89
	AFIIN Plus	89	89	100	94	491	0	63	89
	Flair	71	70	91	79	313	31	132	71
	Flair Plus	73	73	90	80	310	36	115	73
Tamako Market Latest	Vader	89	89	100	94	489	1	62	10
	TextBlob	93	93	100	97	517	0	37	8

Keyword	Library	Accuracy	Precision	Recall	F1	TP	FN	FP	TN
	TextBlob Plus	93	93	100	97	517	0	37	8
	SentiWordNet	85	85	100	92	461	1	83	17
	SentiWordNet Plus	85	85	100	92	459	2	79	22
	AFIIN	91	91	100	95	504	1	52	5
	AFIIN Plus	91	91	100	95	504	1	52	5
	Flair	80	81	99	89	437	5	106	14
	Flair Plus	73	73	90	80	310	36	115	95
Non Non Biyori Latest	Vader	85	85	99	92	800	8	138	23
	TextBlob	89	89	100	94	867	0	102	0
	TextBlob Plus	89	89	100	94	867	0	102	0
	SentiWordNet	77	77	100	87	738	2	221	8
	SentiWordNet Plus	76	77	100	87	739	1	227	2
	AFIIN	86	86	100	93	838	0	131	0
	AFIIN Plus	86	86	100	93	838	0	131	0
	Flair	75	77	91	83	592	62	180	135
	Flair Plus	75	76	91	83	598	56	189	126
Sewayaki Kitsune no Senko- san The Helpful Fox Senko- san Senko- san Latest	Vader	87	87	100	93	490	1	75	2
	TextBlob	85	86	100	92	485	1	82	0
	TextBlob Plus	85	86	100	92	485	1	82	0
	SentiWordNet	83	84	100	91	474	1	93	0
	SentiWordNet Plus	84	84	100	91	473	1	92	2
	AFIIN	89	89	100	94	498	2	64	4
	AFIIN Plus	89	89	100	94	498	2	63	5
	Flair	68	68	99	81	376	5	174	13
	Flair Plus	71	72	93	81	363	26	141	38

Keyword	Library	Accuracy	Precision	Recall	F1	TP	FN	FP	TN
Shoujo Shuumatsu Ryokou Girls' Last Tour Latest	Vader	80	81	98	88	549	11	132	20
	TextBlob	87	88	100	93	619	1	88	4
	TextBlob Plus	87	97	100	93	619	1	89	3
	SentiWordNet	70	72	75	73	300	99	117	196
	SentiWordNet Plus	71	73	77	75	305	92	117	198
	AFIIN	82	82	99	90	571	5	126	10
	AFIIN Plus	81	82	99	90	569	7	126	10
	Flair	63	67	80	73	350	88	173	101
	Flair Plus	65	68	83	75	368	74	195	95

Berdasarkan tabel 4.1 sedangkan data lengkap tersedia di lampiran 2, selama data aktual positif semua *library* akan cenderung positif dengan akurasi 75%-98%, hanya beberapa judul yang memiliki statistik rendah. Selain itu kecenderungan yang terjadi antara lain:

1. VADER selalu stabil bagus dalam statistik.
2. TextBlob selalu memiliki kecenderungan positif meskipun *library* lain dalam suatu judul ke arah negatif karena tidak ada filter *subjectivity*.
3. SentiWordNet memiliki statistik rendah.
4. AFIIN jika datanya ke positif bisa memiliki statistik terbaik, namun saat data tercampur dengan negatif, prediksi cenderung negatif.
5. Flair memiliki statistik terburuk dan sering melakukan kesalahan klasifikasi.
6. Pembersihan tambahan hanya memberikan dampak akurasi sekitar 0% - 3%. Hal ini terjadi karena kata dalam *hashtag* dan *emoji* selalu selaras dengan *tweet*. *Emoji* positif bersama *tweet* positif dan sebaliknya. Sehingga perubahan positif ke negatif dan sebaliknya jarang terjadi kecuali pada *tweet* pendek.
7. Judul yang dianggap negatif karena memiliki kata negatif cenderung merusak akurasi seperti *Usagi Srop*, *Mr. Villain's Day Off*, *Slow Loop*, *My New Boss Is Goofy*, dan *One Off*.

4.1.10 Excel Analisis Sentimen

Pada excel HasilAkurasi_311810002.xlsx pada PKL belum mendukung format untuk digunakan pada *website* karena setiap baris harus terisi sehingga perlu perubahan manual agar bisa digunakan.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
1	Judul	Library	Tweet Total	Accuracy	Precision	Recall	F1-Score	TP	FN	FP	TN	FN	FP	Error									
2		Vader	83	83	98	80	348	7	71	38	7	64	0										
3		TextBlob	89	89	100	94	411	0	53	0	0	0	40										
4		TextBlob Plus	88	89	100	94	410	1	53	0	1	51	0										
5		SentiWordNet	71	73	96	83	318	14	120	12	11	108	1										
6		Mushishi	404	73	74	96	84	322	12	115	15	12	111	2									

Gambar 4.21 Excel HasilAkurasi_311810002.xlsx

Pada file excel, baris dan kolom yang digabung dengan *merge center* akan membuat *error* saat filtering, contoh *Mushishi* VADER akan menjadi Mushishi dengan judul *Mushishi*, akan tetapi *library* lain akan memiliki judul NaN karena data kosong. Oleh karena itu, perbaikan dilakukan secara manual dengan membuat file excel baru dengan mengambil data penting yaitu judul, *library*, *total tweet*, true positive, true negative, persentase positif dengan perhitungan (true positive + False Negative) / Total *Tweet*.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	Judul	Library	Tweet	TP	FN	Persentase Positif										
2	Mushishi	Vader	464	348	7	76,50862069										
3	Mushishi	TextBlob	464	411	0	88,57758621										
4	Mushishi	TextBlob Plus	464	410	1	88,57758621										
5	Mushishi	SentiWordNet	464	318	14	71,55172414										
6	Mushishi	SentiWordNet Plus	464	322	12	71,98275862										
7	Mushishi	AFIN	464	392	0	84,48275862										
8	Mushishi	AFIN Plus	464	391	1	84,48275862										
9	Mushishi	Flair	464	339	6	74,35344828										
10	Mushishi	Flair Plus	464	338	12	75,43103448										
11	Barakamon	Vader	485	366	7	76,90721649										
12	Barakamon	TextBlob	485	431	0	88,86597938										
13	Barakamon	TextBlob Plus	485	431	0	88,86597938										
14	Barakamon	SentiWordNet	485	375	4	78,1443299										
15	Barakamon	SentiWordNet Plus	485	377	2	78,1443299										
16	Barakamon	AFIN	485	403	0	83,09278351										
17	Barakamon	AFIN Plus	485	403	0	83,09278351										
18	Barakamon	Flair	485	343	24	75,67010309										
19	Barakamon	Flair Plus	485	344	26	76,28865979										
20	Natsume Yuujinchou	Vader	601	551	0	91,68053245										
21	Natsume Yuujinchou	TextBlob	601	539	0	89,68386023										
22	Natsume Yuujinchou	TextBlob Plus	601	539	0	89,68386023										
23	Natsume Yuujinchou	SentiWordNet	601	504	2	84,19301165										

Gambar 4.22 Excel sentimen_analisis.xlsx

Pada gambar 4.22, setiap baris sudah terisi dengan data sehingga bisa ditampilkan ke *website* tanpa terjadi *error* karena data kosong.

4.2 Hasil Pengambilan Data

4.2.1 Data Anime dari API Anilist

Data anime berhasil diambil dari Anilist melalui API, menghasilkan 21.345 judul unik yang disimpan dalam file *anime_data.xlsx*. Data mencakup informasi seperti judul romaji, judul Inggris, *link* Anilist, rating skala 1–100, genre, unsur dengan persentasenya, status sumber asli, cerita fiksi atau bukan, apakah ada unsur LGBT, Apakah ada unsur kekerasan berlebihan? Apakah ceritanya berat?, dan *link streaming*.

Data terimpan pada excel *anime_data.xlsx* sebanyak 21.345 judul unik. Berikut contoh datanya:

Tabel 4.2 Data Anime *Shingeki no Kyojin* pada excel *anime_data.xlsx*

Data	Nilai
Romaji Title	Shingeki no Kyojin
English Title	Attack on Titan
Anilist link	https://anilist.co/anime/16498/Attack-on-Titan/
Genres	Action, Drama, Fantasy, Mystery
Tags	Revenge 97%, Kaiju 93%, Male Protagonist 89%, Military 88%, Tragedy 88%, Post-Apocalyptic 84%, Primarily Teen Cast 84%, Super Power 82%, Steampunk 79%, Henshin 79%, Swordplay 76%, Gore 75%, Memory Manipulation 75%, Dystopian 74%, Ensemble Cast 73%, Orphan 73%, Kuudere 70%, Espionage 68%, Cannibalism 66%, Coming of Age 62%, Shounen 61%, Survival 60%, Medieval 60%, Vore 60%, Time Skip 59%, Amnesia 58%, Rural 51%, Suicide 50%, CGI 46%, Primarily Male Cast 44%, Adoption 35%
Rating	84
Original Work Finished	Yes
Fiction	Yes
LGBT	No
Extreme Violence	Yes
Requires Thinking	Yes

Data	Nilai
Streaming Links	Crunchyroll http://www.crunchyroll.com/attack-on-titan , Hulu http://www.hulu.com/attack-on-titan , Crunchyroll https://www.crunchyroll.com/attack-on-titan-dubs

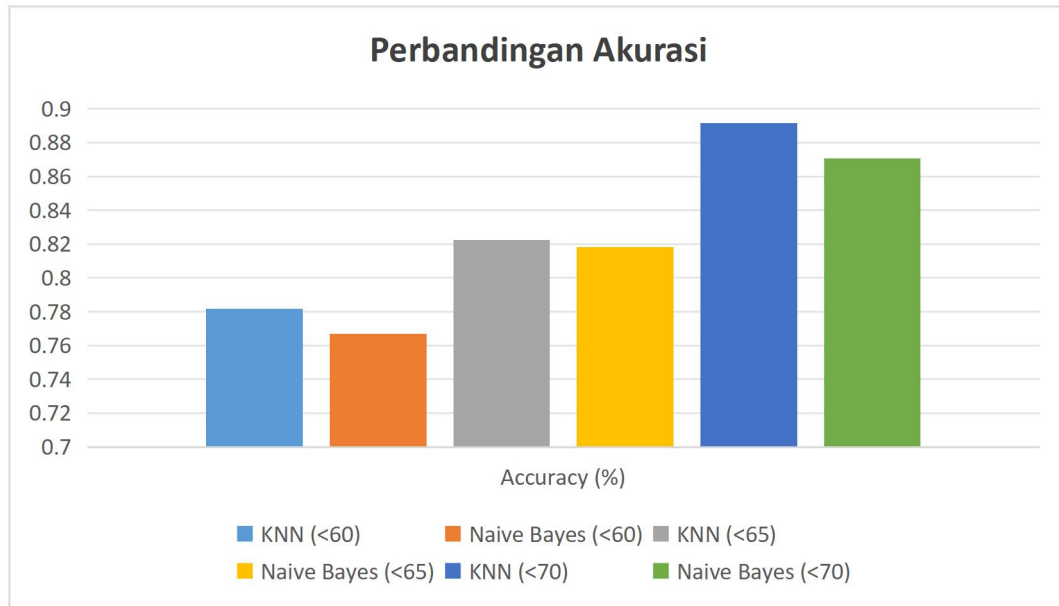
Dari semua data penting yang dibutuhkan untuk website terambil, dan dilakukan beberapa tahap tambahan sesuai yang dijelaskan pada bab 3. *Rating*, *Original Work Finished*, *Fiction*, *LGBT*, *Extreme Violence*, dan *Requires Thinking* digunakan sebagai pertanyaan wajib pada filter *form* dan diurutkan dari rating tertinggi. Semua genre, unsur, dan 6 pertanyaan *form* digunakan untuk perhitungan KNN. *Romaji Title* dan *English Title* untuk pengecekan *fuzzy matching* dan *autocomplete* fitur pencarian anime termirip. *Anilist link* dan *Streaming Links* digunakan pada hasil rekomendasi semua fitur, kecuali tabel analisis sentimen tidak mengeluarkan *Anilist link*.

4.2.2 Pertanyaan Form Filter Anime

Form filter memungkinkan pengguna menyaring anime berdasarkan enam kriteria atau pertanyaan yaitu rating (diambil dari *average score* skala 1–10), status sumber asli berdasarkan status *finished* atau *releasing* dari sumber seperti manga, *light novel*, atau novel. Fiksi / non-fiksi. unsur LGBT+, kekerasan ekstrim, dan cerita berat ditentukan dari adanya genre dan unsur spesifik.

4.3 Perbandingan KNN dan Naive Bayes

Untuk mencari model terbaik dari kedua algoritma, rating anime dipisah menjadi 3 tipe yaitu menjadi <60 jelek, 60-80 bagus, >=80 sangat bagus. Tipe 2 adalah <65 jelek, 65-80 bagus, >=80. Tipe 3 adalah <70 jelek, 70-80 bagus, >=80 sangat bagus, maka confusion matriks dapat dilihat pada grafik berikut:



Gambar 4.23 Perbandingan Akurasi 6 Model Pencarian anime Termirip

Tabel 4.3 Perbandingan Confusion Matrix 6 Model Pencarian Anime Termirip

Metrik	KNN (<60)	Naive Bayes (<60)	KNN (<65)	Naive Bayes (<65)	KNN (<70)	Naive Bayes (<70)
Accuracy (%)	0.7819	0.7671	0.8223	0.8183	0.8914	0.8706
Precision	0.78	0.77	0.81	0.81	0.87	0.88
Recall	0.78	0.77	0.82	0.82	0.89	0.87
F1-Score	0.77	0.76	0.81	0.82	0.87	0.87
Average Similarity	0.5802	0.3510	0.5802	0.3510	0.5733	0.3510
Coverage	0.8955	1	0.8955	1	0.9411	1
Diversity	2,8457	4,0775	2,8457	4,2804	2,8197	4,5063

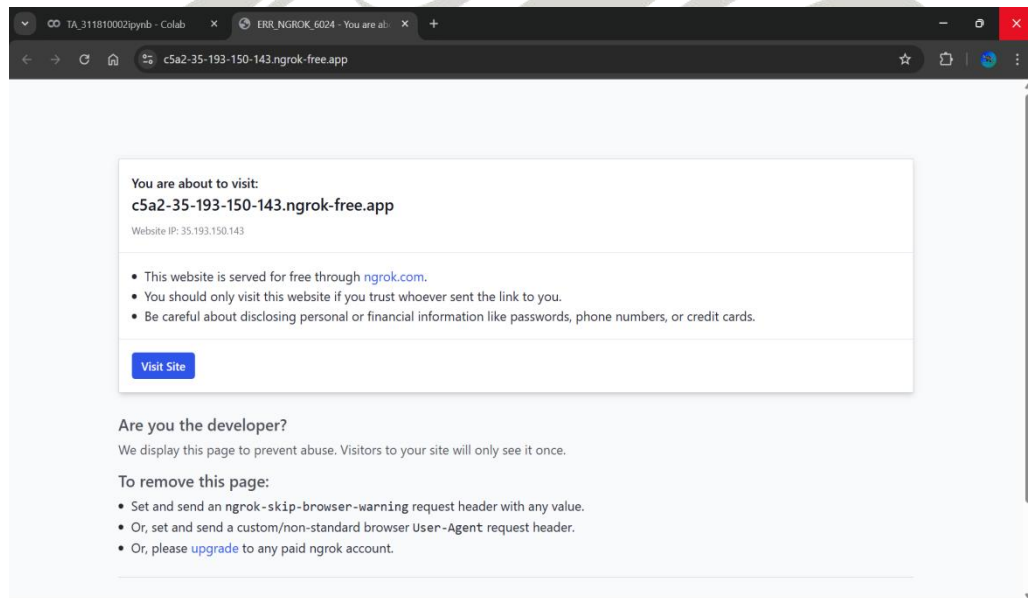
Berdasarkan gambar 4.23 dan tabel 4.3, dipilih KNN dengan pembatasan rating <70 jelek, 70-80 bagus, >=80 sangat bagus menggunakan euclidian distance. KNN akan dipakai pada *website* sebagai algoritma untuk menetapkan anime terbaik berdasarkan genre dan unsur dengan k atau jumlah rekomendasi sebanyak 10 karena 50 rekomendasi terlalu banyak.

4.4 Pengujian Kodinng

Setelah koding dibuat, maka *website* telah berhasil dibuat. File yang dibutuhkan seperti *anime_data.xlsx*, *sentimen_analisis.xlsx*, dan gambar diupload ke *session storage* dan jalankan baris *pip* dan *auth token*.

4.4.1 Link ngrok

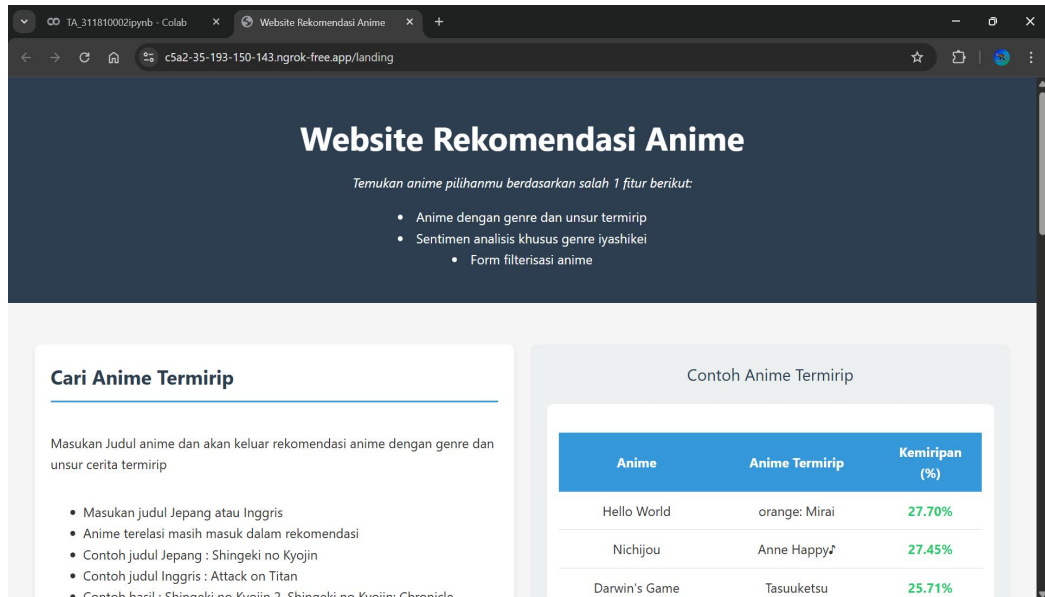
Setelah menjalankan bagian koding dengan fungsi *main*, maka akan keluar *link* ngrok, pastikan sudah mengupload *file* excel dan gambar ke *session storage*. Untuk ke *website* yang dibuat, klik *link* ngrok dan saat pertama kali akan keluar halaman khusus verifikasi sebelum ke *webiste* yang dibuat. Klik *visit site* dan akan menuju *landing page website*.



Gambar 4.24 Halaman *About to Visit* ngrok

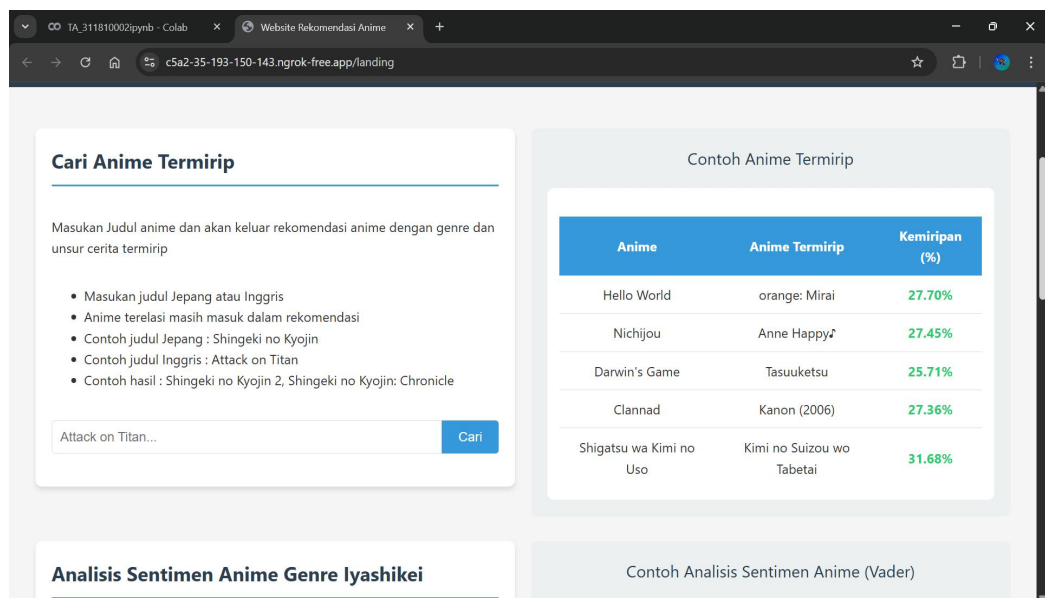
4.4.2 Landing page

Saat memasuki *website*, halaman yang ditampilkan adalah halaman pertama website dalam rancang bangun ini diberi nama */landing* dengan judul bernama *Anime Recommendation System* sesuai dengan nama yang ditulis pada fungsi `<title> </title>`.



Gambar 4.25 Landing page

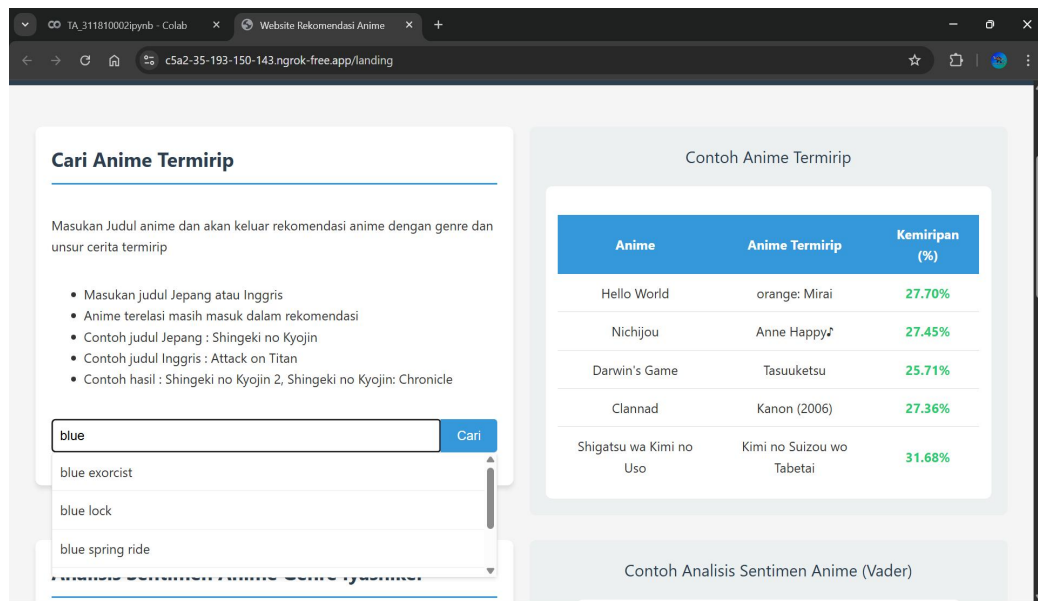
Link ngrok setelah ini akan selalu menuju *landing page* kecuali *main* dijalankan ulang maka akan membuat *link* ngrok baru sehingga jika ada halaman statis yang ingin ditampilkan pada halaman lain harus dikoding manual. Pada bagian *header* berisi nama *website* dan 3 fitur penting yang dapat digunakan.



Gambar 4.26 Fitur Cari Anime Termirip Berdasarkan Genre dan Unsur

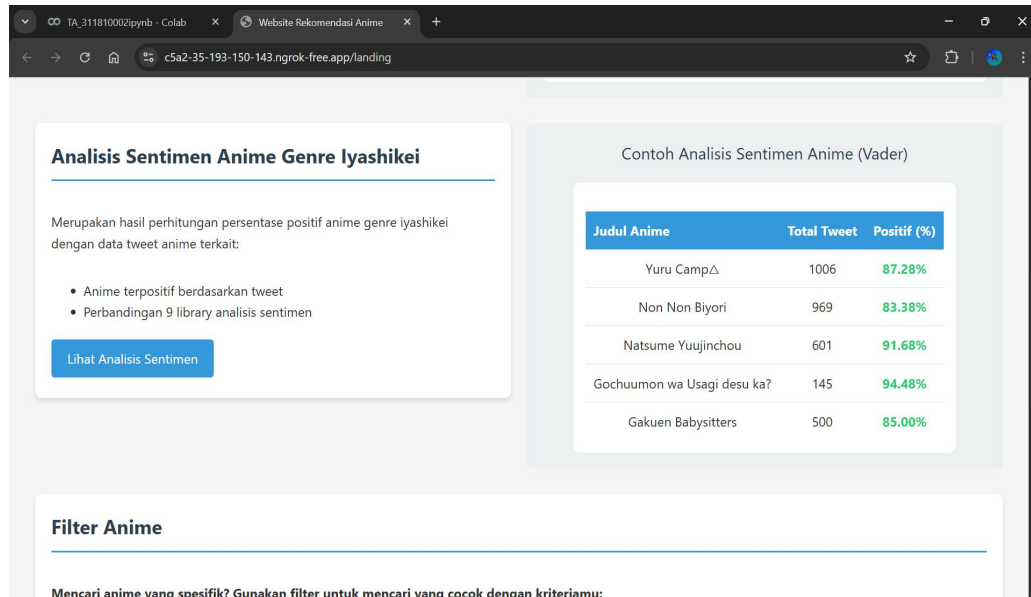
Pada gambar 4.26, fungsi pertama website adalah pencarian anime termirip berdasarkan 6 kriteria, genre, dan unsur menggunakan KNN. Sebelum

search box, sudah diberikan langkah pengisian dan contohnya. Pada bagian kanan terdapat contoh pencarian anime termirip dan persentase kemiripan agar pengguna dapat mendapatkan gambaran hasil saat tombol cari diklik. Saat menulis di *search box*, akan keluar opsi judul anime yang sedang dicari agar memudahkan pengguna jika lupa judul lengkap yang sedang dicari.



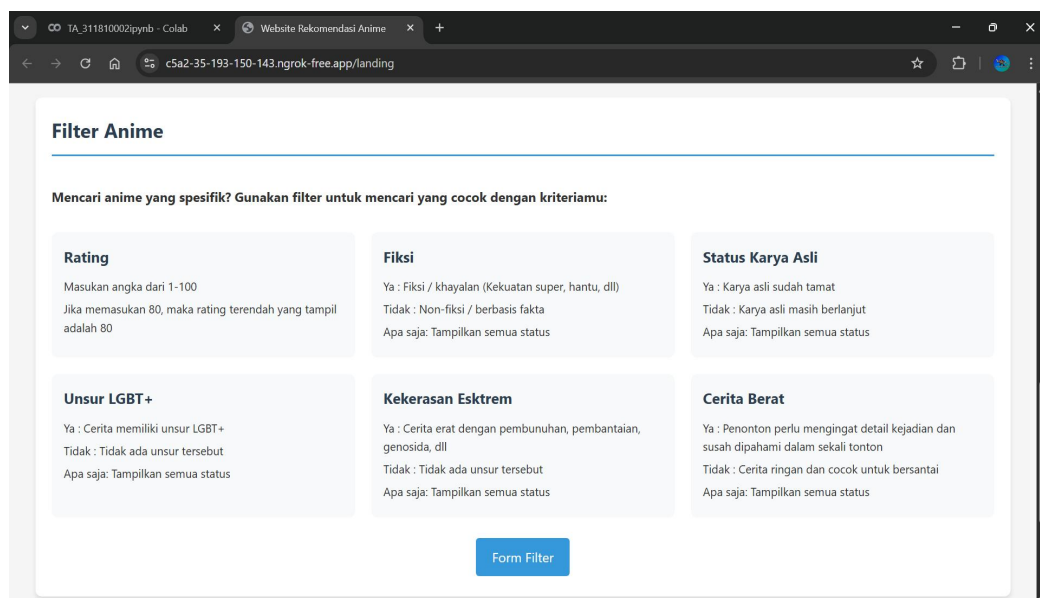
Gambar 4.27 Fitur *autocomplete*

Pada gambar 4.27, saat mengetik *blue* akan keluar 5 rekomendasi dengan sistem *drag down* yaitu *blue exorcist*, *blue lock*, *blue spring ride*, *perfect blue*, dan *grand blue*. Input akan dicocokkan dengan *fuzzy matching* dengan seluruh judul romaji dan judul Inggris dan akan ditampilkan oleh *autocomplete*.



Gambar 4.28 Fitur Analisis Sentimen Anime Genre *Iyashikei*

Pada gambar 4.28, sentimen anime genre *iyashikei* dalam format tabel. Hasil excel analisis_sentimen.xlsx akan difilter per *library* dan perbandingan semua *library* tiap judul agar mudah dibaca dan VADER dijadikan pilihan pada contoh sebelah kanan karena secara hasil VADER merupakan yang terbaik dalam analisis sentimen media sosial.



Gambar 4.29 Penjelasan *Form Filter Anime*

Gambar 4.29 adalah fitur ketiga yaitu filter anime menggunakan *form* dengan diajukan 6 pertanyaan yang wajib dijawab dengan detail pertanyaan dan opsi jawaban yang diambil berdasarkan pertanyaan umum penikmat anime di internet yang dapat dikodingkan yaitu rating, fiksi, status karya asli, unsur LGBT+, kekerasan ekstrim, dan cerita berat. Saat klik tombol *form* filter, akan berpindah halaman pengisian *form*.

segit325@gmail.com'."/>

Gambar 4.30 *Footer*

Bagian *footer* berisi nama *website* dan tahun sebagai *copyright*, nama pembuat, universitas dan logo, serta email pembuat. Pada bagian ini, bisa ditambahkan *link* sosial media lain dan hal umum seperti FAQ.

4.4.3 Hasil Pencarian Anime Termirip Berdasarkan Genre dan Unsur

Setelah memasukan judul pada *search box*, maka *input* akan dicek dengan *fuzzy matching* dengan judul romaji dan judul Inggris di *anime_data.xlsx*.

Hasil "blue lock" **100% match**

Rating	Genre	Unsur	Fiksi	Cerita Asli Tamat	Unsur LGBT+	Kekerasan Ekstrem	Cerita Berat
80.0	Action, Drama, Sports	Football, Male Protagonist, Shounen	Ya	Tidak	Tidak	Ya	Ya

Anime Termirip (Minimal 17.5%)

1) Blue Lock VS. U-20 JAPAN **34.93%**

Judul Inggris: BLUE LOCK Season 2

Rating: 75.0

2) Blue Lock: EPISODE Nagi **33.76%**

Judul Inggris: BLUE LOCK THE MOVIE -EPISODE NAGI-

Rating: 74.0

Gambar 4.31 Rekomendasi Anime Termirip dengan Blue Lock (Bagian 1)

Pada bagian atas akan ditampilkan judul, persentase kecocokan *fuzzy matching*, rating, genre, 3 unsur dengan persentase tertinggi, apakah fiksi, status cerita asli, unsur LGBT+, unsur kekerasan, dan apakah cerita berat. Anime termirip pada saat perancangan akan ditampilkan sebanyak 50 judul yang sudah diatur pada KNN dengan syarat minimal 17.5 kecocokan. Namun, karena 50 terlalu banyak, akhirnya *website* hanya menampilkan maksimal 10 judul.

Anime Termirip (Minimal 17.5%)

1) Blue Lock VS. U-20 JAPAN **34.93%**

Judul Inggris: BLUE LOCK Season 2

Rating: 75.0

Genres: Action, Drama, Sports

Unsur: Football, Male Protagonist, Primarily Male Cast

Streaming: Crunchyroll, Bilibili TV

[Cek Anilist](#)

2) Blue Lock: EPISODE Nagi **33.76%**

Judul Inggris: BLUE LOCK THE MOVIE -EPISODE NAGI-

Rating: 74.0

Genres: Action, Drama, Sports

Unsur: Football, Primarily Male Cast, Male Protagonist

Streaming: Crunchyroll

[Cek Anilist](#)

3) World Trigger 3rd Season **25.60%**

Rating: 82.0

Genres: Action, Sci-Fi

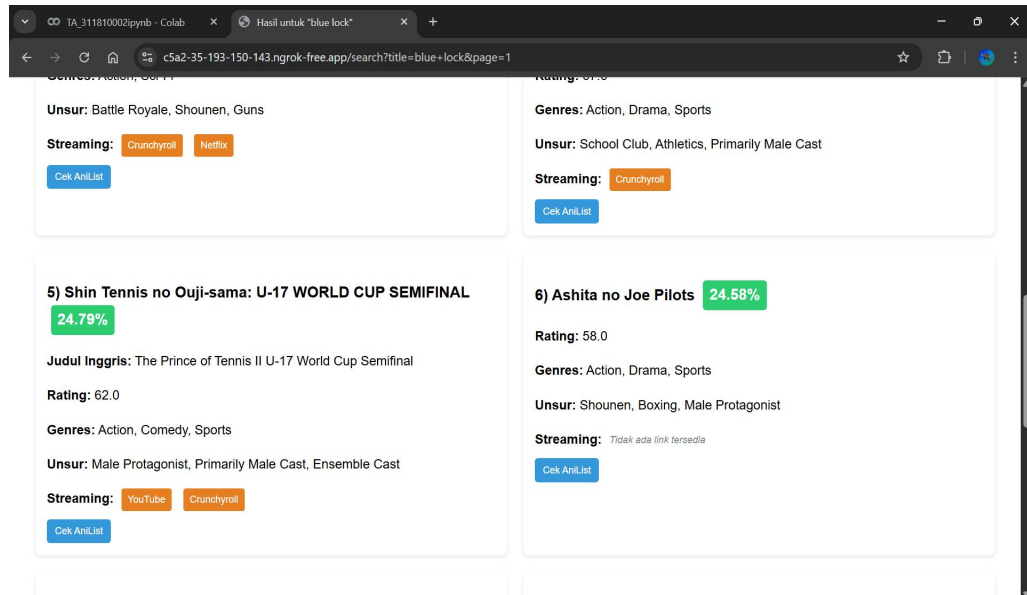
4) Shakunetsu Kabaddi **25.37%**

Judul Inggris: Burning Kabaddi

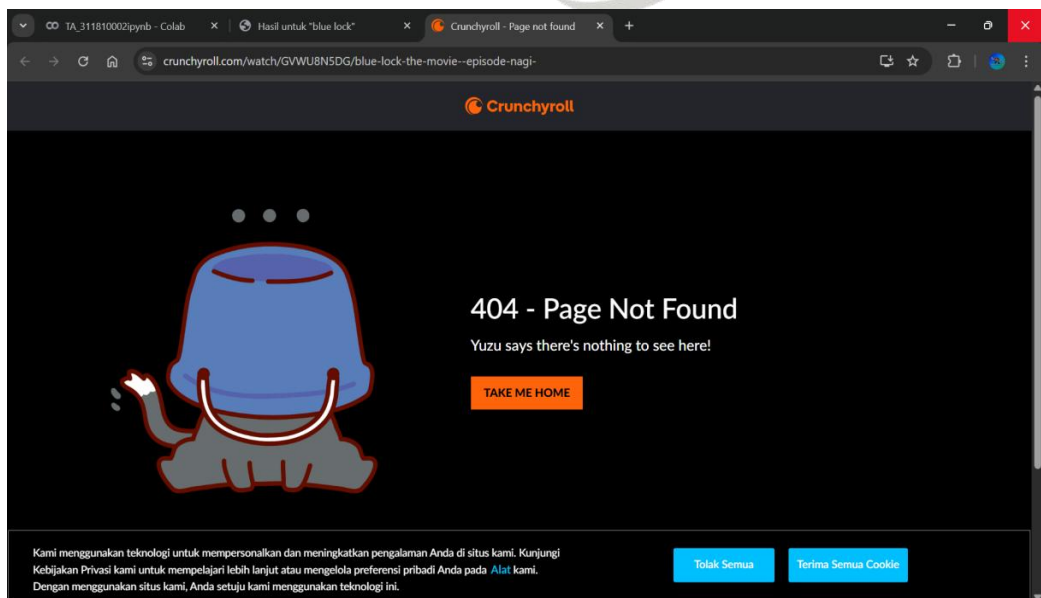
Rating: 67.0

Gambar 4.32 Rekomendasi Anime Termirip dengan Blue Lock (Bagian 2)

Setiap judul akan diberi peringkat, judul, tingkat kecocokan, rating, genre, 3 unsur teratas, *link streaming*, dan *link* ke Anilist. Jika tidak ada *link streaming* akan memunculkan tulisan “tidak ada link tersedia”.



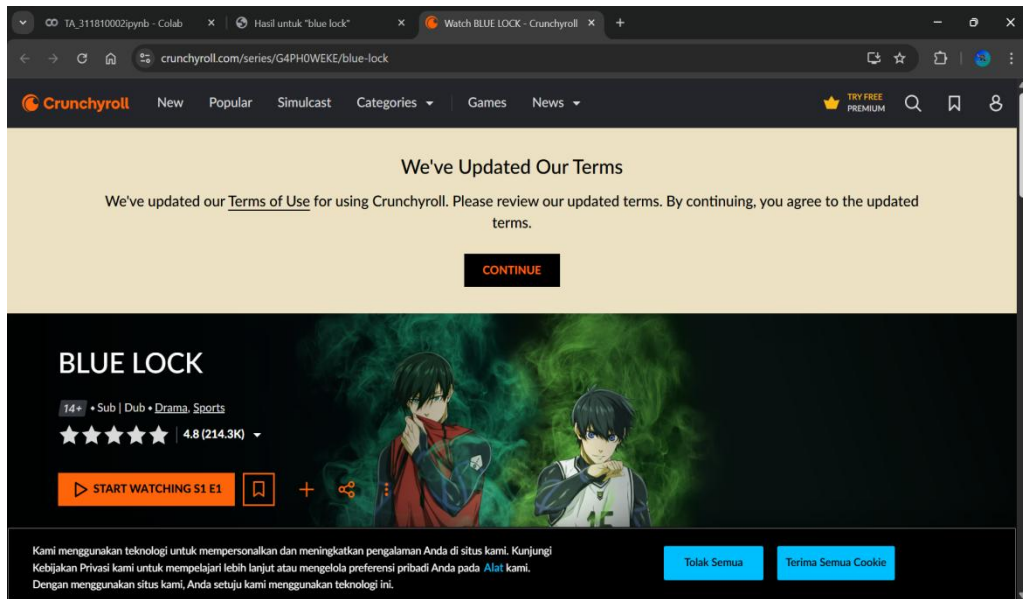
Gambar 4.33 Hasil Rekomendasi Jika Tidak Ada *Link Streaming*



Gambar 4.34 Website Streaming Page Not Found

Sumber: Crunchyroll

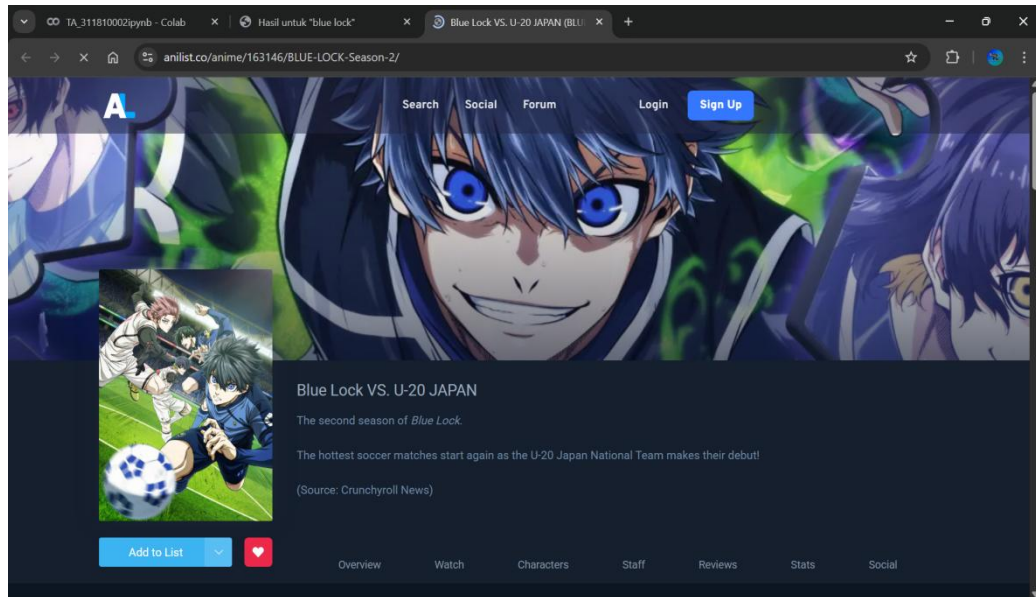
Ada beberapa halaman dihapus pada *website streaming* karena hanya ditayangkan beberapa minggu kemudian dihapus. Pada Anilist, *link website* yang sudah tidak ada videonya masih ada sehingga ikut terambil saat pengambilan data menggunakan API. Data seperti ini tidak bisa dihapus pada excel secara otomatis karena harus dihapus manual.



Gambar 4.35 *Website Streaming* jika ada

Sumber: Crunchyroll

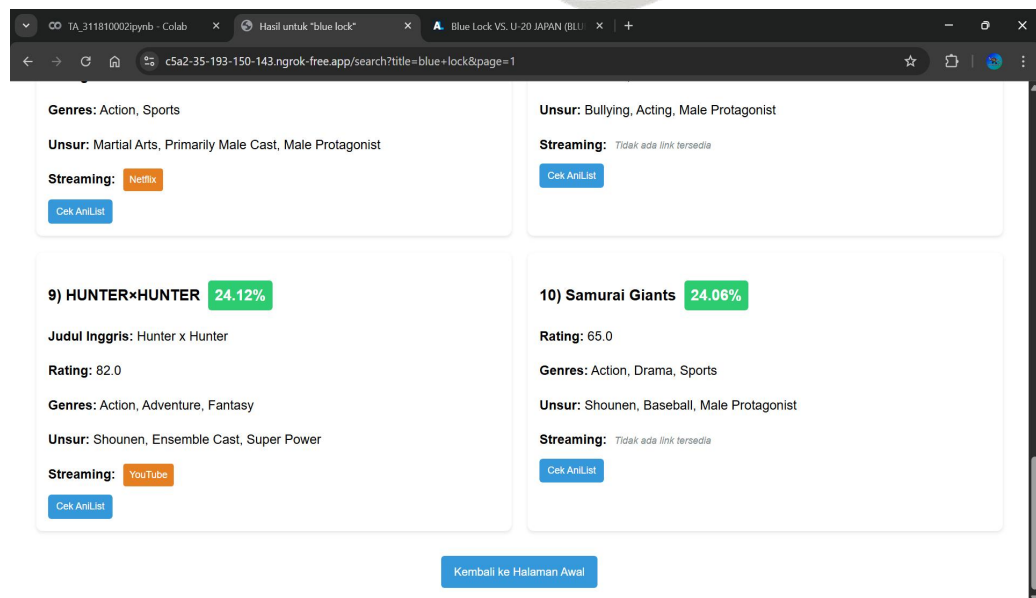
Contoh jika *link streaming* mengarah ke *website* yang masih ada datanya. *Website streaming* akan berguna agar pengguna tidak kesusahan mencari tempat menonton legal setelah merasa cocok dengan rekomendasi.



Gambar 4.36 Website Anilist

Sumber: Anilist

Contoh halaman Anilist untuk Blue Lock VS. U-20 JAPAN. Sangat dibutuhkan bagi beberapa pengguna yang membutuhkan informasi lebih detail.



Gambar 4.37 Tombol ke Halaman Awal

Pada bagian bawah akan terdapat tombol kembali ke halaman awal. Paginasi tidak dibutuhkan karena semua rekomendasi sudah dimunculkan pada 1

halaman saja. Contoh hasil pencarian anime termirip dengan *Blue Lock* yang memiliki genre action, drama, dan sports serta 3 unsur teratas *football*, *male protagonist*, dan *shounen* antara lain:

Tabel 4.4 10 Anime termirip dengan Blue Lock berdasarkan KNN Euclidian

No.	Judul	Kemiripan	Rating	Genre	Unsur
1	Blue Lock VS. U-20 JAPAN	34.93%	75.0	Action Drama Sports	Football Male Protagonist Primarily Male Cast
2	Blue Lock: EPISODE Nagi	33,76%	74.0	Action Drama Sports	Football Primarily Male Cast Male Protagonist
3	World Trigger 3rd Season	25.60%	82.0	Action Sci-Fi	Battle Royale Shounen Guns
4	Shakunetsu Kabaddi	25.37%	67.0	Action Drama Sports	School Club Athletics Primarily Male Cast
5	Shin Tennis no Ouji-sama: U-17 WORLD CUP SEMIFINAL	24.79%	62.0	Action Comedy Sports	Male Protagonist Primarily Male Cast Ensemble Cast
6	Ashita no Joe Pilots	24.58%	58.0	Action Drama Sports	Shounen Boxing Male Protagonist
7	Hanma Baki	24.45%	76.0	Action Sports	Martial Arts Primarily Male Cast Male Protagonist
8	Nyuugaku Youhei	24.14%	60.0	Action Drama	Bullying Acting Male Protagonist
9	HUNTER×HUNTER	24.12%	82.0	Action Adventure Fantasy	Shounen Ensemble Cast Super Power
10	Samurai Giants	24.06%	65.0	Action Drama Sports	Shounen Baseball Male Protagonist

Tabel 4.4 menunjukan 10 anime termirip dengan *Blue Lock*, mayoritas anime bergenre *action, sports, drama*, serta unsur male protagonist dan *primarily teen cast*. Terdapat anime tanpa genre *sports* seperti *World Trigger 3rd Season*, akan tetapi kebanyakan genre dan unsurnya mirip sehingga secara perhitungan KNN euclidian kemiripannya tinggi. Perhitungan ini membuktikan anime yang termirip dengan *sport* belum tentu *sport*.

Contoh hasil pencarian anime termirip dengan *Yuru Camp* yang memiliki genre *comedy* dan *slice of life* serta 3 unsur teratas *outdoor activities, camping*, dan *iyashikei* antara lain:

Tabel 4.5 10 Anime termirip dengan *Yuru Camp* berdasarkan KNN Euclidian

No.	Judul	Kemiripan	Rating	Genre	Unsur
1	Yuru Camp△ SEASON 2	46.87%	84.0	Comedy Slice of Life	Outdoor Activities Iyashikei Cute Girls Doing Cute Things
2	Heya Camp△	33.73%	72.0	Comedy Slice of Life	Cute Girls Doing Cute Things Outdoor Activities Iyashikei
3	Yakunara Mug Cup Mo	31.77%	63.0%	Comedy Slice of Life	Cute Girls Doing Cute Things Primarily Female Cast Iyashikei
4	Yuru Camp△ SEASON 4	31.18%	60.0	Comedy Slice of Life	Outdoor Activities Cute Girls Doing Cute Things Iyashikei
5	Yakunara Mug Cup Mo: Niban Gama	30.68%	74.0	Comedy Slice of Life	Cute Girls Doing Cute Things School Club Female Protagonist
6	Yuru Camp△ SEASON 3	30.68%	80.0	Comedy Slice of Life	Outdoor Activities Iyashikei Cute Girls Doing Cute Things

No.	Judul	Kemiripan	Rating	Genre	Unsur
7	Yuru Camp△ SEASON 3 OVA	29.98%	75.0	Comedy Slice of Life	Iyashikei Camping Female Protagonist
8	Houkago Teibou Nisshi	29.91%	71.0	Comedy Slice of Life Sport	Fishing School Club Cute Girls Doing Cute Things
9	Yama no Susume: Next Summit	29.88%	77.0	Adventure Comedy Slice of Life	Outdoor Activities School Club Primarily Female Cast
10	Yama no Susume	29.74%	67.0	Adventure Comedy Slice of Life	Outdoor Activities Primarily Female Cast Iyashikei

Dari tabel 4.5 menunjukan 10 anime termirip dengan *Yuru Camp*, terdapat kesamaan genre *comedy*, *slice of life* dan unsur *Outdoor Activities*, *Camping*, *Cute Girls Doing Cute Things* dan *iyashikei*. Tipe cerita yang ringan, karakter mayoritas perempuan, dan klub sekolah. Contoh hasil pencarian anime termirip dengan *Hello World* yang memiliki genre *drama*, *psychological*, dan *romance* serta 3 unsur teratas *time manipulation*, *virtual world*, dan *alternate universe* antara lain:

Tabel 4.6 10 Anime termirip dengan *Hello World* berdasarkan KNN Euclidian

No.	Judul	Kemiripan	Rating	Genre	Unsur
1	orange: Mirai	27.70%	73.0	Drama Psychologi cal Romance	Shoujo School Male Protagonist
2	Ashita Sekai ga Owaru toshitemo	26.99%	59.0	Drama Romance Sci-Fi	Alternate Universe Full CGI Magic
3	Toki wo Kakeru Shoujo	26.70%	78.0	Drama Romance Sci-Fi	Time Manipulation Coming of Age Female

No.	Judul	Kemiripan	Rating	Genre	Unsur
					Protagonist
4	Kimi wo Aishita Hitori no Boku e	26.21%	73.0	Romance Sci-Fi	Alternate Universe Male Protagonist Memory Manipulation
5	Sekai ga Horobiru Mae ni Kimi ni Aitai	26.15%	40.0	Drama Romance Sci-Fi	Alternate Universe
6	Natsuiro no Sunadokei	26.12%	60.0	Comedy Drama Romance	Time Manipulation Male Protagonist Female Harem
7	Boku ga Aishita Subete no Kimi e	26.01%	74.0	Romance Sci-Fi	Coming of Age Alternate Universe Male Protagonist
8	Soutai Sekai	25.90%	51.0	Drama Sci-Fi	Full CGI Alternate Universe Idol
9	Jiaonang Jihua	25.85%	60.0	Drama Psychological Sci-Fi	Memory Manipulation
10	Tenchi Muyou! in LOVE 2: Haruka Naru Omoi	25.69%	69.0	Drama Romance Sci-Fi	Shounen Memory Manipulation Alternate Universe

Tabel 4.6 menunjukan 10 anime termirip dengan *Hello World*, terdapat kesamaan genre *drama*, *sci-fi* atau *science fiction* serta genre *alternate universe*, *time manipulation*, dan *memory manipulation*. Tipe cerita yang berat, berkaitan dengan fiksi seperti dunia lain, penghapusan memori, psikologi, dan manipulasi waktu. Contoh ketika mengetik *attack on titan* justru *typo* menjadi *ssysvl pm yoysm* akan keluar halaman *error* karena *typo* yang terlalu parah.



Gambar 4.38 Halaman Pencarian Anime Termirip Jika *Typo* Terlalu Parah

Halaman khusus dimunculkan jika *typo* terlalu parah, akan meminta pengguna memperhatikan penulisan dan tersedia tombol kembali ke halaman awal.

4.4.4 Sentimen Analisis Anime Genre *Iyashikei*

Setelah pengguna menekan tombol lihat analisis sentimen, *website* akan berpindah ke halaman berisi judul, *total tweet*, dan persentase positif dengan *library* VADER sebagai *default*. Pada bagian kanan atas, terdapat pilihan *library* lain yaitu TextBlob, TextBlob plus, SentiWordNet, SentiWordNet Plus, AFIIN, AFIIN plus, Flair, Flair plus, dan bandingkan semua. Tiap halaman menampilkan 5 judul dengan total judul sebanyak 61 untuk pilihan bukan bandingkan semua dan untuk bandingkan 1 judul 1 halaman.

Analisis Sentimen Anime (Vader)

Algoritma: Vader

Judul Anime	Total Tweet	Persentase Positif	Link Streaming
Super Cub	131	98.47%	Crunchyroll
Sketchbook ~full color's~	86	97.67%	Crunchyroll
Yotsuuro Biyori	86	97.67%	Crunchyroll
Cool Doji Danshi	165	96.36%	Crunchyroll YouTube Bilibili TV Netflix
Danchigai	95	94.74%	Crunchyroll

Awal Sebelum 1 2 3 4 5 6 7 8 9 10 ... Sesudah Akhir

Kembali ke Halaman Awal

Gambar 4.39 Perbandingan Anime Terpositif *Library* VADER

Analisis Sentimen Anime (TextBlob)

Algoritma: TextBlob

Judul Anime	Total Tweet	Persentase Positif	Link Streaming
Yoru no Kuni	122	97.54%	Netflix
My New Boss Is Goofy	144	97.22%	Crunchyroll
Cool Doji Danshi	165	96.36%	Crunchyroll YouTube Bilibili TV Netflix
Super Cub	131	96.18%	Crunchyroll
Sketchbook ~full color's~	86	95.35%	Crunchyroll

Awal Sebelum 1 2 3 4 5 6 7 8 9 10 ... Sesudah Akhir

Kembali ke Halaman Awal

Gambar 4.40 Perbandingan Anime Terpositif *Library* TextBlob

Analisis Sentimen Anime (TextBlob Plus)

Algoritma: TextBlob Plus

Judul Anime	Total Tweet	Persentase Positif	Link Streaming
Yoru no Kuni	122	97.54%	Netflix
My New Boss Is Goofy	144	97.22%	Crunchyroll
Cool Doji Danshi	165	96.36%	Crunchyroll YouTube Bilibili TV Netflix
Super Cub	131	96.18%	Crunchyroll
Sketchbook ~full color's~	86	95.35%	Crunchyroll

Awal Sebelum 1 2 3 4 5 6 7 8 9 10 ... Sesudah Akhir

Kembali ke Halaman Awal

Gambar 4.41 Perbandingan Anime Terpositif *Library* TextBlob Plus

Analisis Sentimen Anime (SentiWordNet)

Algoritma: SentiWordNet

Judul Anime	Total Tweet	Persentase Positif	Link Streaming
Emiya-san Chi no Kyou no Gohan	132	93.18%	Crunchyroll
So Ra No Wo To	142	92.96%	Crunchyroll Bilibili TV Netflix
Yoru no Kuni	122	92.62%	Netflix
Maiko-san Chi no Makanai-san	162	91.98%	Crunchyroll Netflix Bilibili TV Hulu
Sketchbook ~full color's~	86	91.86%	Crunchyroll

Awal Sebelum 1 2 3 4 5 6 7 8 9 10 ... Sesudah Akhir

Kembali ke Halaman Awal

Gambar 4.42 Perbandingan Anime Terpositif *Library* SentiWordNet

Analisis Sentimen Anime (SentiWordNet Plus)

Algoritma: SentiWordNet Plus

Judul Anime	Total Tweet	Persentase Positif	Link Streaming
Emiya-san Chi no Kyou no Gohan	132	93.94%	Crunchyroll
So Ra No Wo To	142	92.96%	Crunchyroll, Bilibili TV, Netflix
Yoru no Kuni	122	92.62%	Netflix
Maiko-san Chi no Makanai-san	162	91.98%	Crunchyroll, Netflix, Bilibili TV, Hulu
Sketchbook ~full color's~	86	91.86%	Crunchyroll

Awal Sebelum 1 2 3 4 5 6 7 8 9 10 ... Sesudah Akhir

Kembali ke Halaman Awal

Gambar 4.43 Perbandingan Anime Terpositif *Library* SentiWordNet Plus

Analisis Sentimen Anime (AFIIN)

Algoritma: AFIIN

Judul Anime	Total Tweet	Persentase Positif	Link Streaming
Super Cub	131	97.71%	Crunchyroll
Sketchbook ~full color's~	86	96.51%	Crunchyroll
Yotsuuro Biyori	86	96.51%	Crunchyroll
Yoru no Kuni	122	95.90%	Netflix
Danchigai	95	95.79%	Crunchyroll

Awal Sebelum 1 2 3 4 5 6 7 8 9 10 ... Sesudah Akhir

Kembali ke Halaman Awal

Gambar 4.44 Perbandingan Anime Terpositif *Library* AFIIN

Analisis Sentimen Anime (AFIIN Plus)

Algoritma: AFIIN Plus

Judul Anime	Total Tweet	Persentase Positif	Link Streaming
Super Cub	131	97.71%	Crunchyroll
Sketchbook ~full color's~	86	96.51%	Crunchyroll
Yotsuiri Biyori	86	96.51%	Crunchyroll
Yoru no Kuni	122	95.90%	Netflix
Danchigai	95	95.79%	Crunchyroll

Awal Sebelum 1 2 3 4 5 6 7 8 9 10 ... Sesudah Akhir

Kembali ke Halaman Awal

Gambar 4.45 Perbandingan Anime Terpositif *Library* AFIIN Plus

Analisis Sentimen Anime (Flair)

Algoritma: Flair

Judul Anime	Total Tweet	Persentase Positif	Link Streaming
Sketchbook ~full color's~	86	91.86%	Crunchyroll
Yotsuiri Biyori	86	91.86%	Crunchyroll
Mahoraba: Heartful Days	120	90.00%	Crunchyroll
Yoru no Kuni	122	86.89%	Netflix
Kami-tachi ni Hirowareta Otoko	145	86.21%	Hulu YouTube Netflix Crunchyroll

Awal Sebelum 1 2 3 4 5 6 7 8 9 10 ... Sesudah Akhir

Kembali ke Halaman Awal

Gambar 4.46 Perbandingan Anime Terpositif *Library* Flair

Analisis Sentimen Anime (Flair Plus)

Algoritma: Flair Plus

Judul Anime	Total Tweet	Persentase Positif	Link Streaming
Yoru no Kuni	122	90.16%	Netflix
Mahoraba: Heartful Days	120	90.00%	Crunchyroll
Cool Doji Danshi	165	86.67%	Crunchyroll YouTube Bilibili TV Netflix
Mitsuboshi Colors	59	86.44%	Tidak ada link tersedia
Amaama to Inazuma	606	85.64%	Bilibili TV

Awal Sebelum 1 2 3 4 5 6 7 8 9 10 ... Sesudah Akhir

Kembali ke Halaman Awal

Gambar 4.47 Perbandingan Anime Terpositif *Library* Plus

Super Cub

Total Tweet: 131

Streaming: Crunchyroll

Library	Persentase Positif
Vader	98.47%
AFIIN	97.71%
AFIIN Plus	97.71%
TextBlob	96.18%
TextBlob Plus	96.18%
Flair	83.97%
Flair Plus	83.21%
SentiWordNet	68.70%
SentiWordNet Plus	67.94%

Gambar 4.48 Perbandingan 9 *Library* Terhadap *Super Cub*

4.4.5 Form Filter anime

Pengguna akan memilih opsi jawaban dari 6 pertanyaan wajib. Jika terdapat pilihan kosong tidak akan berpindah ke halaman hasil dan memunculkan peringatan. Terdapat tombol filter anime dan tombol kembali ke halaman awal.

Filter Anime

Minimal Rating (1-100): *

Apakah cerita fiksi? *

Cerita asli sudah tamat? *

Mengandung unsur LGBT+? *

Terdapat kekerasan ekstrem? *

Gambar 4.49 Halaman Form Anime dan Peringatan

Jika sudah diisi lengkap, maka akan berpindah ke halaman hasil.

Kriteria pilihanmu :

Rating Minimal	Fiksi	Cerita Asli Tamat	Unsur LGBT+	Kekerasan Ekstrem	Cerita Berat
80	Tidak	Ya	Tidak	Tidak	Tidak

Anime Tercocok

1) Haikyuu!!: Karasuno Koukou VS Shiratorizawa Gakuen Koukou

Rating: 87.0

Genres: Comedy, Drama, Sports

Unsur: Volleyball, School Club, Shounen

Streaming: [Crunchyroll](#) [Hulu](#) [Netflix](#) [YouTube](#) [Bilibili TV](#)

[Cek AniList](#)

2) THE FIRST SLAM DUNK

Rating: 87.0

Genres: Sports

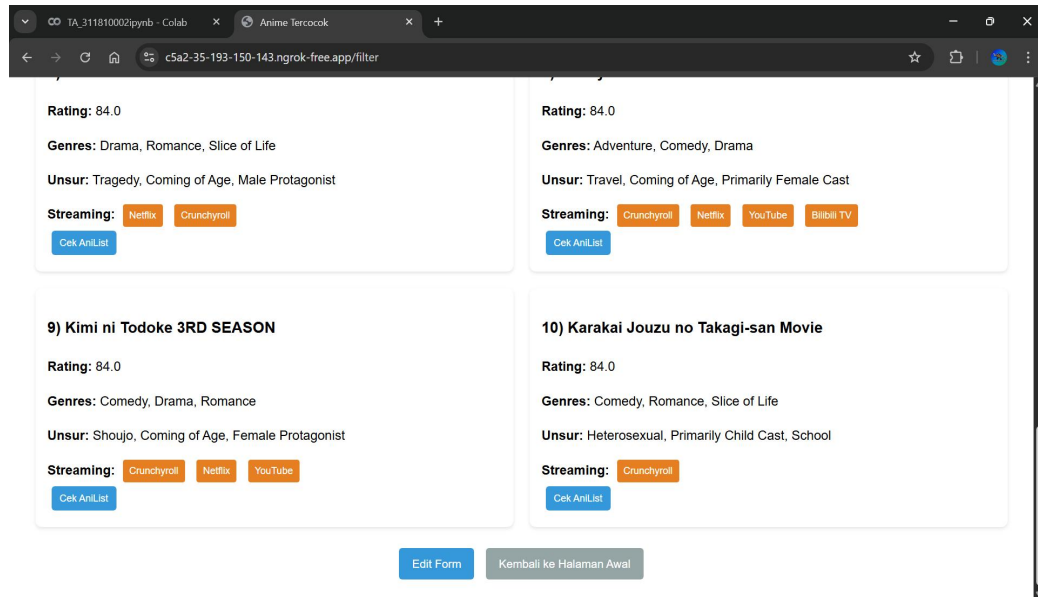
Unsur: Basketball, School Club, Coming of Age

Streaming: [Crunchyroll](#) [Bilibili TV](#)

[Cek AniList](#)

Gambar 4.50 Halaman Filter *Form* Anime

Bagian awal halaman akan menunjukkan kriteria yang dipilih saat mengisi form. Anime tercocok diurutkan dari rating tertinggi berisi judul, rating, genre, 3 unsur tertinggi, *link streaming*, dan *link Anilist*.



Gambar 4.51 Tombol Edit *Form* dan Kembali ke Halaman Awal

Pada bagian bawah halaman terdapat tombol paginasi, tombol *edit form* untuk mengembalikan pengguna ke halaman pengisi *form* yang terisi kriteria terakhir, dan tombol kembali ke halaman awal. Tidak perlu paginasi karena semua rekomendasi sudah ditampilkan pada 1 halaman saja.

Contoh hasil dengan kriteria rating minimal 80, non-fiksi, cerita asli tamat, tidak ada unsur LGBT+, tidak ada unsur kekerasan, dan cerita tidak berat adalah sebagai berikut:

Tabel 4.7 10 Anime teratas dengan kriteria rating minimal 80, non-fiksi, cerita asli tamat, tidak ada unsur LGBT+, tidak ada unsur kekerasan, dan cerita tidak berat

No.	Judul	Rating	Genre	Unsur
1	Haikyuu!! Karasuno VS Shiratorizawa Gakuen Koukou	87.0	Comedy Drama Sports	Volleyball School Club Shounen
2	THE FIRST SLAM DUNK	87.0	Sports	Basketball School Club Coming of Age
3	Haikyuu!! 2nd Season	86.0	Comedy Drama Sports	Volleyball School Club Ensemble Cast
4	Haikyuu!! TO THE TOP 2	86.0	Comedy Drama	Volleyball Male Protagonist

No.	Judul	Rating	Genre	Unsur
			Sports	Primarily Male Cast
5	Chihayafuru 3	85.0	Drama Romance Slice of Life	Shoujo Coming of Age Female Protagonist
6	Shigatsu wa Kimi no Uso	84.0	Drama Music Romance	Tragedy Primarily Teen Cast Disability
7	Kimi no Suizou wo Tabetai	84.0	Drama Romance Slice of Life	Tragedy Coming of Age Male Protagonist
8	Sora yori mo Tooi Basho	84.0	Adventure Comedy Drama	Travel Coming of Age Primarily Female Cast
9	Kimi ni Todoke 3RD SEASON	84.0	Comedy Drama Romance	Shoujo Coming of Age Female Protagonist
10	Karakai Jouzu no Takagisan Movie	84.0	Comedy Romance Slice of Life	Heterosexual Primarily Child Cast School

Tabel 4.7 menunjukan 10 data teratas yang memenuhi kriteria, anime yang memenuhi kriteria di atas kebanyakan adalah *sports*, *romance*, cerita sedih, dan musik.

Sementara itu jika difilter dengan kriteria rating minimal 80, fiksi, cerita asli tamat, tidak ada unsur LGBT+, ada unsur kekerasan, dan cerita berat adalah sebagai berikut:

Tabel 4.8 10 Anime teratas dengan kriteria rating minimal 80, fiksi, cerita asli tamat, tidak ada unsur LGBT+, ada unsur kekerasan, dan cerita berat

No.	Judul	Rating	Genre	Unsur
1	Gintama: THE FINAL	91.0	Action Comedy Drama	Samurai Swordplay Dissociative Identities
2	Hagane no Renkinjutsushi: FULLMETAL	90.0	Action Adventure Drama	Alchemy Conspiracy Military

No.	Judul	Rating	Genre	Unsur
	ALCHEMIST			
3	Gintama°	90.0	Action Comedy Drama	Samurai Tragedy Meta
4	Shingeki no Kyojin 3 Part 2	89.0	Action Drama Fantasy	Dystopian Military Conspiracy
5	Steins;Gate	89.0	Drama, Psychologic al Sci-Fi	Time Manipulation Alternate Universe Time Loop
6	Fruits Basket: The Final	89.0	Comedy Drama Psychologic al	Shoujo Found Family Coming of Age
7	Gintama!: Enchousen	89.0	Action Comedy Drama	Parody Shounen Found Family
8	Mob Psycho 100 II	88.0	Action Comedy Drama	Super Power Coming of Age Urban Fantasy
9	MONSTER	88.0	Drama Horror Mystery	Crime Seinen Detective
10	BLEACH: Sennen Kessen-hen	88.0	Action Adventure Supernatural	War Swordplay Shounen

Tabel 4.8 menunjukkan 10 data teratas. Anime yang memenuhi kriteria di atas kebanyakan adalah *action*, *adventure*, *war*, *criminal*, *tragedy*, dan *conspiracy*.

Sementara itu jika difilter dengan kriteria rating minimal 80, fiksi maupun non fiksi, cerita asli belum tamat, tidak ada unsur LGBT+, tidak ada unsur kekerasan, dan cerita berat maupun ringan adalah sebagai berikut:

Tabel 4.9 10 Anime teratas kriteria rating minimal 80, fiksi maupun non fiksi, cerita asli belum tamat, tidak ada unsur LGBT+, tidak ada unsur kekerasan, dan cerita berat maupun ringan

No.	Judul	Rating	Genre	Unsur
1	Sousou no Frieren	91.0	Adventure Drama Fantasy	Travel Magic Elf

No.	Judul	Rating	Genre	Unsur
2	Boku no Kokoro no Yabai Yatsu 2nd Season	87.0	Comedy Romance Slice of Life	School Heterosexual Male Protagonist
3	Haikyuu!!: Gomi Suteba no Kessen	86.0	Comedy Drama Sports	Volleyball Shounen School Club
4	NANA	85.0	Drama Music Romance	Female Protagonist Shoujo Primarily Adult Cast
5	Hajime no Ippo: New Challenger	85.0	Comedy Drama Sports	Boxing Shounen Primarily Male Cast
6	Hajime no Ippo: Rising	85.0	Comedy Drama Sports	Boxing Shounen Primarily Male Cast
7	Natsume Yuujinchou Roku	85.0	Drama Fantasy Slice of Life	Youkai Iyashikei Male Protagonist
8	Dandadan	84.0	Action Comedy Drama	Urban Fantasy Youkai Ghost
9	Yuru Camp△ SEASON 2	84.0	Comedy Slice of Life	Outdoor Activities Iyashikei Cute Girls Doing Cute Things
10	Natsume Yuujinchou Go	84.0	Drama Fantasy Slice of Life	Iyashikei Youkai Male Protagonist

Tabel 4.9 menunjukan 10 data teratas yang memenuhi kriteria. Anime yang memenuhi kriteria kebanyakan adalah cerita tanpa kekerasan sehingga genre *adventure* dan *action* ringan masih termasuk. Namun ada kondisi tidak ada rekomendasi karena tidak ada yang sesuai kriteria, Contoh kriteria rating minimal 90, non fiksi, cerita asli tamat, tidak ada unsur LGBT+, tidak ada unsur kekerasan, dan cerita tidak berat.

Kriteria pilihanmu :					
Rating Minimal	Fiksi	Cerita Asli Tamat	Unsur LGBT+	Kekerasan Ekstrim	Cerita Berat
90	Tidak	Ya	Tidak	Tidak	Tidak

Tidak ada anime yang cocok dengan kriteriamu.

[Edit Form](#) [Kembali ke Halaman Awal](#)

Gambar 4.52 Halaman Filter *Form* Jika Tidak Ada yang Memenuhi Kriteria

Pada percobaan ini, jika rating terlalu tinggi tidak ada anime yang direkomendasi dengan pesan tidak anime yang cocok dengan kriteriamu.

4.5 Diskusi

Tugas akhir yang dilaksanakan dalam beberapa bulan ini menghasilkan koding yang mampu mengeluarkan rekomendasi anime termirip berdasarkan 6 kriteria, genre dan unsur, analisis sentimen anime genre iyashikei, dan form filterisasi anime dengan baik. Berikut adalah diskusi mengenai hasil percobaan.

1. Secara umum, hasil confusion matrix sudah bagus untuk sentimen analisis, akurasi 75% - 98% menunjukkan model bisa memprediksi teks dengan aktual. Akan tetapi, confusion matrix belum mempunyai fungsi bawaan untuk hanya memperlihatkan jumlah dan persentase data aktual positif. Pada website ini, dibuat ke excel baru agar bisa menampilkan jumlah *tweet* dan persentase positif.
2. Pembersihan data untuk analisis sentimen masih belum maksimal. Penghilangan *username* dan *link* mirip *stopword* karena mereka tidak memiliki skor maka diskip saat pencocokan ke *library* sehingga kurang memberi efek. *Hastag* yang dihapus masih menyimpan kata yang

tergabung sehingga tidak cocok dengan *library* sehingga perlu ada koding khusus agar kata bisa terpisah sempurna. Fungsi mengubah *emoji* menjadi kata dapat bekerja dengan baik.

3. Hasil confusion matriks dengan 3 macam pengelompokan rating menunjukkan KNN lebih unggul dari Naive Bayes. Pemisahan rating <60 jelek, 60-80 bagus, >=80 KNN memiliki Accuracy 78.19% sementara Naive Bayes 76,71%. Pemisahan rating <65 jelek, 65-80 bagus, >=80 sangat bagus, KNN memiliki Accuracy 82.23% sementara Naive Bayes : 81.83%. Pemisahan rating <70 jelek, 70-80 bagus, >=80 sangat bagus, KNN memiliki Accuracy 89.14% sementara Naive Bayes : 87.06%.
4. Anime yang direkomendasikan menggunakan KNN dengan jarak euclidian dengan 10-Fold *cross-validation* memiliki Accuracy: 89.14%, Precision: 0.87, Recall: 0.89, F1-Score: 0.87, *Average Similarity*: 0.5733, *Coverage*: 0.9411, dan *Diversity*: 2.8197. Berdasarkan hasil pada tabel 4.5, 4.6, dan 4.7, anime yang direkomendasikan memiliki unsur dan tema yang mirip namun bervariasi secara rating bahkan anime dengan rating 58% tetap direkomendasikan.
5. *Form* dapat berjalan lancar dengan pengajuan 6 pertanyaan dengan opsi jawab telah didetailkan pada landing page. Jika tidak sesuai rekomendasi maka tidak ada yang ditampilkan.
6. *Website* tidak mengalami *bug* dan rata-rata perpindahan halaman setelah *input* data sekitar 3-8 detik seperti pencarian anime termirip yang terhitung normal untuk *website prototype*. Jika hanya berpindah halaman seperti mengklik *form* filter dan mundur ke halaman awal membutuhkan waktu 1 detik. Program dibuat di Google Colab dengan data excel diupload ke *session storage* bisa dipercepat prosesnya dengan optimalisasi lebih lanjut.

Koding masih bisa dikembangkan lebih lanjut karena masih ada beberapa keterbatasan seperti disebutkan di bawah ini:

1. Keyword Twitter

Beberapa anime jika menggunakan data *latest* sangat sedikit sehingga harus merubah *tab latest* ke *top* akan tetapi jumlahnya hanya 100-200 *tweet*. Jika data kurang banyak maka tidak perlu spesifikasi *keyword* seperti anime dengan permasalahan akan tercampur dengan penelurusan lain atau penghapuan teks secara manual. Belum ada dalam pembahasan jurnal yang membahas minimal data optimal untuk analisis sentimen maupun SVC, akan tetapi dalam hasil koding tidak ditemukan overfitting dan underfitting walaupun beberapa akurasi kurang.

2. Jumlah data *harvest*

Meskipun memakai *keyword* yang sama, terkadang jumlah *tweet* bisa berubah drastis. Tidak ada penjelasan pasti kenapa bisa terjadi.

3. Pembersihan kata analisis sentimen

Pembersihan kata memberikan perubahan kecil antara 0% - 3%, sebagian besar justru mengurangi daripada meningkatkan akurasi. Pembersihan lain yang masih dilakukan antara lain pembenaran *typo* dan penghilangan bahasa bukan Inggris pada beberapa *tweet*.

4. Sentiment positif, neutral, dan negatif

Textblob memiliki fungsi jika *subjectivity* 0 atau fakta, nilai *polarity* akan 0. Fungsi yang tidak digunakan ini menyebabkan semua hasil akurasi Textblob cenderung ke positif. Ada baiknya jika hasil analisis tetap dalam positif, netral, dan negatif, atau bisa lebih dari 3 kelas. Akan tetapi kekurangan banyaknya pelabelan adalah mempersusah prediksi di confusion matrix.

5. *Library* analisis sentimen

Kasus yang ditemukan adalah kata yang memiliki multi makna sering disalahartikan serta frasa yang bersifat subjek atau objek seharusnya tidak dikalkulasi, Kasus dalam penelitian adalah *Attack on Titan* yang merupakan judul diskor negatif karena pada *library*, *attack* memiliki skor negatif. *Cool Doji Danshi* diberi skor positif karena *cool* dianggap positif.

6. KNN

Pada pembuatan website ini mengambil seluruh unsur yang akan digunakan sebagai kriteria *pada* form dan perhitungan KNN dengan jarak

euclidian. Sebaiknya dalam pembuatan rekomendasi sejenis sebaiknya menggunakan beberapa filter data misalnya hanya mengambil unsur dengan persentase >50% dan mengambil semua unsur. Kedua hasil filter bisa dites untuk melihat model mana yang lebih benar.

7. Filterisasi tambahan

Pada anime termirip menggunakan genre dan unsur memiliki hasil yang variatif sehingga beberapa pengguna akan heran kenapa anime *sports* bisa diberikan rekomendasi anime tanpa genre *sports* seperti pada tabel 4.5 ketika *Blue Lock* mirip dengan *World Trigger 3rd Season*. Oleh karena itu, perlu adanya fitur tambahan setelah filter pertama keluar yaitu adanya unsur dan genre yang wajib ada dari anime yang diketik pada *search box*. Semisal jika user memasukan *Blue Lock*, maka akan keluar pilihan semua unsur dan genre *Blue Lock*. Setelah itu, pengguna bisa memilih semisal wajib ada *sports* sehingga rekomendasi setelah filter kedua hanya menyisakan anime yang memiliki genre *sports*. Hal seperti ini akan diperlukan semisal pengguna sangat butuh unsur atau genre yang sangat spesifik.

8. Form Anime

Pertanyaan pada *form* hanya 6 dan dirasa sudah menjawab keinginan pecinta anime, akan tetapi masih bisa ditambahkan namun dibuat tidak wajib dijawab seperti studio dan tahun rilis dengan syarat data bisa didapat melalui API serta bisa dikoding.

9. UI

Secara penampilan sudah rapi dan terbaca jelas dengan panduan penggunaan, akan tetapi masih polos seperti belum ada gambar anime secara spesifik terutama pada hasil rekomendasi. Pengguna yang melihat *website* mungkin tidak akan terlalu tertarik karena kurangnya penggunaan media seperti gambar, *running text*, video, dan sebagainya. Namun penggunaan media tentu memerlukan penyimpanan yang terstruktur dan kapasitas besar.

10. Database

Dalam skala besar, website membutuhkan database seperti MySQL atau SQLite agar dapat menyimpan data yang banyak dan penggabungan beberapa data sekaligus. Jika data dalam excel hanya sedikit, tidak masalah jika diupload ke *session storage* Google Colab asalkan jumlah data yang diupload sampai *wesbite* 100% berjalan memiliki kapasitas yang cukup dan tidak membebani laptop atau pc.

11. Versi mobile

Koding dibuat khusus untuk *website* dan secara tata letak akan berantakan jika diakses pada *handphone*.



UNIVERSITAS
MA CHUNG

BAB V

PENUTUP

5.1 Kesimpulan

Simpulan yang dapat diambil dari pelaksanaan Tugas Akhir Rancang Bangun Website Rekomendasi Anime Berbasis Kecerdasan Buatan adalah sebagai berikut:

1. *Website* dapat berjalan 100% tanpa *bug* sesuai dengan fitur yang ditentukan sebelumnya yaitu pencarian anime termirip berdasarkan genre dan unsur menggunakan KNN, rekomendasi anime genre *iyashikei* berdasarkan analisis sentimen tweet, dan *form* filterisasi anime.
2. Perpindahan halaman tanpa perpindahan data hanya 1 detik, namun jika ada pemrosesan data seperti filter form dan pencarian anime termirip menggunakan KNN menjadi 3 - 8 detik.
3. *Website* berhasil memberikan rekomendasi anime menggunakan form maupun perhitungan KNN terhadap 21.345 judul anime yang diambil menggunakan API Anilist dan analisis sentimen terhadap 61 judul anime genre *iyashikei*.
4. KNN 10-Fold *cross-validation* memiliki Accuracy: 89.14%, Precision: 0.87, Recall: 0.89, F1-Score: 0.87, *Average Similarity*: 0.5733, *Coverage*: 0.9411, dan *Diversity*: 2.8197
5. Rekomendasi anime berdasarkan kemiripan genre dan unsur berhasil menampilkan anime termirip dan berbeda dengan sistem rekomendasi yang sudah ada di mana genre atau unsur utama tidak selalu ada seperti genre *sports* harus direkomendasikan dengan *sports* lain.
6. KNN memiliki confusion matriks yang lebih baik daripada Naive Bayes sehingga sistem rekomendasi berdasarkan kemiripan genre dan unsur memberikan 10 rekomendasi judul yang sangat mirip dan beberapa anime kurang terkenal termasuk dalam rekomendasi. Jika *typo* terlalu parah sudah disiapkan halaman bertuliskan "Anime tidak ditemukan Periksa penulisan, cek typo".

7. Fitur form filterisasi anime dapat bekerja dengan baik dengan diajukan 6 pertanyaan yang wajib dijawab dengan hasil maksimal 10 rekomendasi. Apabila tidak terisi terdapat peringatan jika jawaban wajib diisi. Jika tidak ada anime yang sesuai perimataan pengguna akan menampilkan pesan “Tidak ada anime yang cocok dengan kriteriamu”.
8. Rekomendasi genre *iyashikei* ditampilkan menggunakan tabel dengan akurasi antara 75% - 98% tanpa overfitting dan underfitting dalam tiap fold. VADER merupakan *library* terbaik untuk analisis sentimen karena stabil dalam hal statistik terutama akurasi. Textblob meskipun selalu tinggi akurasi hal ini disebabkan perhitungan dilakukan tanpa filter *subjectivity* sehingga menciptakan banyak positif sehingga hasilnya cukup bias. AFINN bekerja cenderung mengikuti data, tinggi statistik terutama akurasi jika positif dan sebaliknya. Sentiwordnet dan Flair kurang cocok dibuktikan dengan statistik yang cenderung rendah.

5.2 Saran

1. Analisis sentimen bisa menggunakan *scraping* atau *harvest* pada media sosial lain karena *limit* 1000 *scrool* per 15 menit dari Twitter membuat pengambilan data menjadi lambat.
2. Perlu adanya tes untuk negatif, netral, dan positif terutama dalam *keyword* yang memiliki tercampur data netral seperti pengumuman dan data statistik.
3. Menggunakan model dan algoritma klasifikasi lain agar bisa menemukan statistik yang lebih baik seperti model BERT atau Algoritma CNN.
4. Data yang diambil sebaiknya terbaru karena adanya perbandingan dengan anime baru atau perilisan episode terbaru. *Tweet* yang terlalu lama kemungkinan tidak relevan.
5. Pengambilan rekomendasi bisa saja dengan pendetailan seperti tanpa kata seperti *announcement* atau mengambil dari Textblob dengan filter *subjectivity*.
6. Penggunaan database untuk akses data dan penggunaan media seperti foto serta video untuk memperbagus UI.

DAFTAR PUSTAKA

- Ahmed, W., 2024, 'Using Twitter as a data source: An overview of current social media research tools', *VoxPol*, <<https://voxpol.eu/using-twitter-as-a-data-source-an-overview-of-current-social-media-research-tools/>> [Accessed 16 September 2024].
- Akbik, A., Rasul, K., Bergmann, T., Schweter, S., Blythe, D., & Vollgraf, R., 2019, 'FLAIR: An easy-to-use framework for state-of-the-art NLP', *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Demonstrations*, 54–59, <<https://aclanthology.org/N19-4010.pdf>> [Accessed 11 December 2024].
- Alodokter, 2022, 'Cara mengatasi depresi dengan terapi pola pikir dan perilaku', *Alodokter*, <<https://www.alodokter.com/cara-mengatasi-depresi-dengan-terapi-pola-pikir-dan-perilaku>> [diakses pada 9 September 2024].
- Baccianella, S., Esuli, A., & Sebastiani, F., 2010, 'SENTIWORDNET 3.0: An enhanced lexical resource for sentiment analysis and opinion mining', *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*, 2200–2204, <<http://nmis.isti.cnr.it/sebastiani/Publications/LREC10.pdf>> [Accessed 2 Desember 2024].
- Backlinko, 2024, 'Social media usage & growth statistics', *Backlinko*, <<https://backlinko.com/social-media-users/>> [Accessed 10 September 2024].
- Baihaqi, I. R., 2021, 'Penerapan content-based filtering pada sistem rekomendasi anime berbasis semantik', *Repository Wicida*, 2:1–12, <<https://repository.wicida.ac.id/3899/2/1743011-S1-Jurnal.pdf>> [Diakses pada 1 Desember 2024].
- Dixon, S. J., 2024, 'X/Twitter: Global audience 2024', *Statista*, <<https://www.statista.com/statistics/242606/number-of-active-twitter-users-in-selected-countries/>> [Accessed 9 September 2024].
- Dyouri, A., & Walia, A. S., 2024, 'How to build a Flask Python web application from scratch', *DigitalOcean*, <<https://www.digitalocean.com/community/tutorials/how-to-make-a-web-application-using-flask-in-python-3>> [Accessed 4 February 2025].
- Engati, n.d., 'Rule-based system', *Engati*, <<https://www.engati.com/glossary/rule-based-system>> [Accessed 3 February 2025].

- Esuli, A., 2011, 'SentiWordNet: A publicly available lexical resource for opinion mining', *GitHub*, <<https://github.com/aesuli/SentiWordNet>> [Accessed 2 December 2024].
- GeeksforGeeks, 2025, 'How to calculate Jaccard similarity in Python', *GeeksforGeeks*, <<https://www.geeksforgeeks.org/data-science/how-to-calculate-jaccard-similarity-in-python/>> [Accessed 3 February 2025].
- GeeksforGeeks, 2024, 'Differences between black box testing and white box testing', *GeeksforGeeks*, <<https://www.geeksforgeeks.org/differences-between-black-box-testing-vs-white-box-testing/>> [Accessed 18 March 2025].
- GeeksforGeeks, 2024, 'Applying multinomial Naive Bayes to NLP problems', *GeeksforGeeks*, <<https://www.geeksforgeeks.org/applying-multinomial-naive-bayes-to-nlp-problems/>> [Accessed 26 March 2025].
- Goodwin, H., 2024, 'What is an API (application programming interface)?', *IBM*, <<https://www.ibm.com/topics/api>> [Accessed 1 December 2024].
- Hetler, A., 2023, 'What is Twitter?', *TechTarget*, <<https://www.techtarget.com/whatis/definition/Twitter>> [Accessed 9 September 2024].
- Hrikesa, I. G. N. A. A., Pratama, I. P. A. E., & Raharja, I. M. S., 2024, 'Sentimen analisis masyarakat terhadap manga One Piece di media sosial Twitter dengan metode SVM', *JITTER: Jurnal Ilmiah Teknologi dan Komputer*, 5:1-8, <<https://ojs.unud.ac.id/index.php/jitter/article/view/115019/55060>> [diakses pada 1 Desember 2024].
- Hutto, C. J., 2014, 'VADER sentiment analysis', *GitHub*, <<https://github.com/cjhutto/vaderSentiment>> [Accessed 2 December 2024].
- IBM, 2023, 'What are support vector machines (SVMs)?', *IBM*, <<https://www.ibm.com/topics/support-vector-machine>> [Accessed 16 September 2024].
- Indrayuni, E., 2018, 'Komparasi algoritma Naïve Bayes dan Support Vector Machine untuk analisa sentimen review film', *Prosiding Seminar Nasional Teknologi Informasi dan Komunikasi*, 14:1–7, AMIK BSI Pontianak, Pontianak, Indonesia, <<https://ejournal.nusamandiri.ac.id/index.php/pilar/article/view/36/31>> [diakses pada 16 September 2024].

- Koehrsen, W., & Urwin, M., 2024, 'Precision and recall: How to evaluate your classification model', *Built In*, <<https://builtin.com/data-science/precision-and-recall>> [Accessed 16 September 2024].
- Loria, S., 2021, 'TextBlob: Simplified text processing', *GitHub*, <<https://github.com/sloria/TextBlob>> [Accessed 2 Desember 2024].
- Lowm, P., 2024, 'Social media sentiment analysis: Benefits and guide for 2024', *Sprout Social*, <<https://sproutsocial.com/insights/social-media-sentiment-analysis/>> [Accessed 16 September 2024].
- Mailo, F.F, & Lazuradi, L., 2019, "Analisis Sentimen Data Twitter Menggunakan Metode Text Mining Tentang Masalah Obesitas di Indonesia," *J. Inf. Syst. Public Heal.*, vol. 4, no. 1, pp. 28–36, 2019. <<https://jurnal.ugm.ac.id/jisph/article/view/44455/30078>>, [diakses pada 16 September 2024].
- Malde, R., 2020, 'A short introduction to VADER', *Towards Data Science*, <<https://medium.com/data-science/an-short-introduction-to-vader-3f3860208d53>> [Accessed 16 September 2024].
- Malik, U., 2023, 'Python for NLP: Introduction to the Pattern library', *Stack Abuse*, <<https://stackabuse.com/python-for-nlp-introduction-to-the-pattern-library/>> [Accessed 16 September 2024.]
- Mayur, 2018, 'How KNN uses distance measures?', *Analytics Vidhya*, <<https://www.analyticsvidhya.com/blog/2021/08/how-knn-uses-distance-measures/>> [Accessed 3 February 2025].
- Miesle, P., 2023, 'What is cosine similarity: A comprehensive guide', *DataStax*, <<https://www.datastax.com/guides/what-is-cosine-similarity>> [Accessed 3 February 2025].
- MyAnimeList, 2024, 'Top anime', *MyAnimeList*, <<https://myanimelist.net/topanime.php?limit=26450>> [Accessed 9 September 2024].
- MyAnimeList, 2024, 'Iyashikei anime', *MyAnimeList*, <<https://myanimelist.net/anime/genre/63/Iyashikei>> [Accessed 9 September 2024].
- Ngrok, n.d., 'Ngrok: Putting your APIs online shouldn't be hard', *Ngrok*, <<https://ngrok.com/>> [Accessed 3 February 2025].

- Nielsen, F., 2011, 'AFINN: A simple and efficient sentiment analysis tool for the Danish language', *GitHub*, <<https://github.com/fnielsen/afinn>> [Accessed 2 Desember 2024].
- Perry, P. O., 2011, 'AFINN sentiment lexicon', *Corpus Text*, <https://corpus-text.com/reference/sentiment_afinn.html> [Accessed 16 September 2024].
- Plaosan, S. V., n.d., 'Support Vector Machine (SVM)', *LearningBox*, <<https://learningbox.coffeecup.com/SVM.html>> [Accessed 9 September 2024].
- Salsabila, N. A., 2023, 'Analisis sentimen pada media sosial Twitter terhadap tokoh Gus Dur menggunakan metode Naïve Bayes dan Support Vector Machine (SVM)', *Jurnal Fakultas Sains dan Teknologi UIN Syarif Hidayatullah Jakarta*, <<https://repository.uinjkt.ac.id/dspace/handle/123456789/66253>> [diakses pada 1 Desember 2024].
- Satria, H., 2024, 'Cara crawl/mendapatkan data Twitter dengan filter waktu dan lainnya', *Helmi Satria*, <<https://helimisatria.com/blog/cara-crawl-mendapatkan-data-twitter-dengan-filter-waktu-dan-lainnya/>> [diakses pada 1 Desember 2024].
- Scikit-learn, 2025, 'Naive Bayes', *Scikit-learn*, <https://scikit-learn.org/stable/modules/naive_bayes.html> [Accessed 25 March 2025].
- Shah, P., 2020, 'Sentiment analysis using TextBlob', *Towards Data Science*, <<https://towardsdatascience.com/my-absolute-go-to-for-sentimen-analysis-textblob-3ac3a11d524>> [Accessed 16 September 2024].
- Sharma, A., 2024, 'Cross validation in machine learning', *GeeksforGeeks*, <<https://www.geeksforgeeks.org/cross-validation-machine-learning/>> [Accessed 1 December 2024].
- Shiksha Online, 2023, 'Text mining in data mining', *Shiksha Online*, <<https://www.shiksha.com/online-courses/articles/text-mining-in-data-mining/>> [Accessed 1 December 2024].
- Sitanggang, A., Harahap, A. D., Karimullah, A., Dewantara, Y. A., & Rozikin, C., 2023, 'Sistem rekomendasi anime menggunakan metode Singular Value Decomposition (SVD) dan cosine similarity', *Jurnal Teknologi Informasi*, 2(2):90–94, <<http://jurnal.utu.ac.id/JTI/article/viewFile/7787/4290>> [diakses pada 1 Desember 2024].

Srivastava, T., 2025, 'Guide to K-Nearest Neighbors algorithm in machine learning', *Analytics Vidhya*, <<https://www.analyticsvidhya.com/blog/2018/03/introduction-k-neighbours-algorithm-clustering/>> [Accessed 3 February 2025].

X Corp, 2024, 'About the X API', *X Developer*, <<https://developer.x.com/en/docs/x-api/getting-started/about-x-api>> [Accessed 1 December 2024].

X Corp, 2024b, 'Rate limits', *X Developer*, <<https://developer.x.com/en/docs/x-api/rate-limits>> [Accessed 1 December 2024]



UNIVERSITAS
MA CHUNG

LAMPIRAN

1. Hasil Analisis Sentimen

A) VADER

True Positive Example:

Text: @claire_yuri_alt Yeah it's not like usagi drop or smth lol you can kindasmell the incest at a glance. It's almost like the ppl who were surprised cofinofandy and leyley had incest there's a certain vibe you learn to pick up on that isn't always canonized but often is.

Actual Label: Positive

Predicted Label: Positive

False Negative Example:

Text: @SoullessSenseii I play Usagi drop or Unwelcome school and try to matchthe purple to the beat. I feel it boosts chances when I do that!

Actual Label: Positive

Predicted Label: Negative

False Positive Example:

Text: Usagi Drop ending discussion in 2024? Ahh these old bones rememberthose days well. If the series was building up to it that be fine still weird for somebut we see it coming. The fact it came out of nowhere like a speeding car is soass.

Actual Label: Negative

Predicted Label: Positive

True Negative Example:

Text: @Blueseven264 @ChibiReviews No clue never finished it. Just rememberher name Being latina. If it pulls a usagai drop ending.

Imma just pretend it never happened just like I did for usagi drop
that ending never happend.

Actual Label: Negative

Predicted Label: Negative

B) TextBlob

True Positive Example:

Text: OH GOD NO NO!!! NOOOOOOOOOOOOOOOOOOOOOOOOOOOO!

ANOTHER USAGI DROP NOOOOOOOOOOOOOOOOOOOOOOOOOO!!!! WHY!? THAT STARTED
SO GREAT MAN WHY!? <https://t.co/4BANJvNNuw>

Actual Label: Positive

Predicted Label: Positive

False Negative Example:

Text: Usagi Drop really rotted the minds of mangaka's for years to
come. Actual Label: Positive

Predicted Label: Negative

False Positive Example:

Text: @raorapanthera And then there s the third option: reading
the mangathat

ruins everything. I m looking at YOU Usagi Drop

Actual Label: Negative

Predicted Label: Positive

True Negative Example:

Text: Its just a typical toxic romance webtoon not Usagi Drop!

Actual Label: Negative

Predicted Label: Negative

100

C) TextBlob + Pembersihan Data

True Positive Example:

Text: Usagi drop bro keep dropping his preferred age dawggg

Actual Label: Positive

Predicted Label: Positive

False Negative Example:

Text: boogiepop was warawanai ergo proxy id invaded bungou stray
dogs flyingwitch usagi drop ore monogatari natsume yuujinchou

Actual Label: Positive

Predicted Label: Negative

False Positive Example:

Text: And then there s the third option: reading the manga that
ruins everything. I m looking at YOU Usagi Drop

Actual Label: Negative

Predicted Label: Positive

True Negative Example:

Text: Usagi drop would have been better without all the nasty ass
lack of seggs>:(

Actual Label: Negative

Predicted Label: Negative

D) SentiWordNet

True Positive Example:

Text: @Kane0Lore Hyouka Kino's Journey (2003) Land of the Lustrous
MoribitoUsagi Drop (Thank God) Shirobako Space Brothers

Actual Label: Positive

Predicted Label: Positive

101

False Negative Example:

Text: @20pesos_Sopa Usagi Drop lol

Actual Label: Positive

Predicted Label: Negative

False Positive Example:

Text: @RicardoRam63404 @TopGyaru In TWO? I knew about Usagi drop
part 2but witch is the other one?

Actual Label: Negative

Predicted Label: Positive

True Negative Example:

Text: @raorapanthera I think Usagi Drop is definitive proof that
the secondisworse

Actual Label: Negative

Predicted Label: Negative

E) SentiWordNet + Pembersihan Data

True Positive Example:

Text: apparently there are a lot of japanese media where a guy
adopts a girl andthey end up getting married. usagi drop was
another one. seriously disgusting.

Actual Label: Positive

Predicted Label: Positive

False Negative Example:

Text: « I hate you » is basic. « I hope the end of your favourite manga gets theUsagi Drop treatment » is terrifying. It s possible. It could be real.

Actual Label: Positive

Predicted Label: Negative

False Positive Example:

Text: 91 Days Astro Boy Bleach Cowboy Bebop Death Parade Excel SagaFate:Zero Guilty Gear Haikyuu! Idolmaster Jigokuraku Kill la Kill Lucky Star MyHero Academia Odd Taxi Psycho-Pass Rosario x Vampire S-Cry-ed Tokyo Ghoul Usagi Drop Wecolme to the ballroom Yu-Gi-Oh (+GX 5Ds ZeXaL)

Actual Label: Negative

Predicted Label: Positive

True Negative Example:

Text: Being a reluctant fan of Oshi no Ko means making I hope they dont usagi drop us jokes and dreading your life as we approach the oncoming train reck that is the anime catching up with the manga ;-;

Actual Label: Negative

Predicted Label: Negative

F) AFINN

True Positive Example:

Text: @Nijigakupriv It came out while I was busy af lol. Still catching upwiththat banished hero anime Close enough welcome back Usagi drop

Actual Label: Positive

Predicted Label: Positive

False Negative Example:

Text: @KillTrevon Wholesome: Usagi Drop(just don't read the mangal Sci-Fi

action: Noein Thoughtful thriller: Paranoia Agent Funny into Sad: Angel BeatsAction that made me feel something: Wonder Egg Priority

Actual Label: Positive

Predicted Label: Negative

False Positive Example:

Text: @chlrfrmgrl They're from Usagi Drop. Daikichi goes to his grandfather's funeral and learns Gpa had an illegitimate daughter. So he raises his own aunt from the age of 6 to adulthood. When she's an adult she decides to marry him (turns out it's legal). The anime only went up to the time skip

Actual Label: Negative

Predicted Label: Positive

True Negative Example:

Text: Usagi Drop We Wuz Kangz edition <https://t.co/1DOrANh4MZ>

Actual Label: Negative

Predicted Label: Negative

G) AFINN + Pembersihan Data

True Positive Example:

Text: ok so true. actors songs connection is a music anime its peak. theres adachi to shimamura...yuri. aharen san wa hakaranai is a comedy. amagi brilliant parkis a comedy. kakushigoto is a comedy. sakamoto desu ga is a comedy. usagi drop is a cute family story

Actual Label: Positive

Predicted Label: Positive

False Negative Example:

Text: Wholesome: Usagi Drop (just don't read the manga) Sci-Fi action: Noein Thoughtful thriller: Paranoia Agent Funny into Sad: Angel Beats Action that made me feel something: Wonder Egg Priority

Actual Label: Positive

Predicted Label: Negative

False Positive Example:

Text: This didn't usagi drop us it usagi piledrove us

Actual Label: Negative

Predicted Label: Positive

True Negative Example:

Text: Usagi Drop F that first half is so wholesome and then ...
Actual Label: Negative
Predicted Label: Negative

H) Flair

True Positive Example:

Text: GM/GN. Usagi Drop is now on @KreatePlatform
(<https://t.co/BqMt9gkT3P>)
- 1/1 E 50 A - thank you and have a nice weekend!
<https://t.co/6Qny6YIN4J>
Actual Label: Positive
Predicted Label: Positive

False Negative Example:

Text: started usagi drop and wow it s cute. Actual Label: Positive
Predicted Label: Negative

False Positive Example:

Text: @EmaLioncourt usagi drop ou bunny drop
Actual Label: Negative
Predicted Label: Positive

True Negative Example:

Text: @kyoruyeye That anime was top tier... But just like Usagi
Drop mangakinda ruined it
Actual Label: Negative
Predicted Label: Negative

I) Flair + Pembersihan Data

True Positive Example:

Text: Usagi drop started of wholesome and look how that ended. So
don t attack ppl for being apprehensive with a weird scenario and
title like this showhas it sunderstandable and there quite a few
other manga I could mention as well. I gladto know this one doesn
t go that way. Actual Label: Positive
Predicted Label: Positive

False Negative Example:

Text: bro do you like wholesome family stories? you're gonna love usagi drop
 Actual Label: Positive
 Predicted Label: Negative

False Positive Example:

Text: Whoever was the translator for Yen Press s Bunny Drop/Usagi Drop I knowit s been a decade but are you ok? I can t imagine going into this series thinkingit s a cute slice of life family drama and then translating the later volumes. TheTRAUMA I would have had
 Actual Label: Negative
 Predicted Label: Positive

True Negative Example:

Text: Have you read usagi drop
 Actual Label: Negative
 Predicted Label: Negative

2. Tabel Perbandingan hasil dari 61 judul anime genre iyashikei

Keyword	Library	Accuracy	Precision	Recall	F1	TP	FN	FP	TN
Mushishi Latest	Vader	83	83	98	80	348	7	71	38
	TextBlob	89	89	100	94	411	0	53	0
	TextBlob Plus	88	89	100	94	410	1	53	0
	SentiWordNet	71	73	96	83	318	14	120	12
	SentiWordNet Plus	73	74	96	84	322	12	115	15
	AFIIN	85	85	100	92	392	0	70	2
	AFIIN Plus	85	85	100	92	391	1	70	2
	Flair	78	78	98	87	339	6	98	21
	Flair Plus	77	78	97	86	338	12	94	20
Barakamon Latest	Vader	84	84	98	90	366	7	72	40
	TextBlob	89	89	100	94	431	0	54	0
	TextBlob Plus	89	89	100	94	431	0	54	0
	SentiWordNet	80	80	99	89	375	4	92	14
	SentiWordNet	80	80	99	89	377	2	93	13

Keyword	Library	Accuracy	Precision	Recall	F1	TP	FN	FP	TN
	rdNet Plus								
	AFIIN	85	85	100	92	403	0	71	11
	AFIIN Plus	85	85	100	92	403	0	71	11
	Flair	83	85	93	89	343	24	60	58
	Flair Plus	83	85	93	89	344	26	61	54
Natsume Yuujinchou Natsume's Book of Friends Latest	Vader	93	93	100	96	551	0	42	8
	TextBlob	91	91	100	95	539	0	54	8
	TextBlob Plus	91	91	100	95	539	0	54	8
	SentiWordNet	85	85	100	92	504	2	87	8
	SentiWordNet Plus	85	85	99	92	502	4	86	9
	AFIIN	91	91	100	95	540	0	52	9
	AFIIN Plus	91	91	100	95	538	2	53	8
	Flair	82	82	100	90	491	0	110	0
	Flair Plus	81	82	99	89	484	7	109	1
Yuru Camp Laid Back Camp Latest	Vader	88	88	100	94	887	1	115	13
	TextBlob	90	90	100	95	896	2	102	6
	TextBlob Plus	89	89	100	94	898	0	108	0
	SentiWordNet	74	77	92	84	664	56	202	84
	SentiWordNet Plus	77	77	99	87	744	10	220	32
	AFIIN	90	91	100	95	900	2	94	10
	AFIIN Plus	90	90	100	95	898	4	95	9
	Flair	76	76	100	86	759	2	240	5
	Flair Plus	77	77	100	87	765	1	234	6
Usagi Drop Latest	Vader	69	56	59	57	110	75	86	250
	TextBlob	78	79	98	87	385	9	105	22
	TextBlob Plus	79	79	98	87	387	7	105	22
	SentiWordNet	71	72	94	82	340	22	130	29

Keyword	Library	Accuracy	Precision	Recall	F1	TP	FN	FP	TN
	SentiWordNet Plus	71	73	93	81	333	26	125	37
	AFIIN	71	59	54	56	97	84	66	274
	AFIIN Plus	71	64	36	46	65	116	37	303
	Flair	72	51	20	28	19	119	27	346
	Flair Plus	73	51	15	23	21	122	20	358
Tanaka-kun wa Itsumo Kedarugoe Tanaka-kun is Always Listless Latest	Vader	87	87	100	93	480	1	71	87
	TextBlob	88	89	100	94	489	1	63	88
	TextBlob Plus	88	89	100	94	489	0	65	88
	SentiWordNet	85	85	100	92	469	2	82	85
	SentiWordNet Plus	85	85	100	92	469	2	82	85
	AFIIN	89	89	100	94	491	0	63	89
	AFIIN Plus	89	89	100	94	491	0	63	89
	Flair	71	70	91	79	313	31	132	71
	Flair Plus	73	73	90	80	310	36	115	73
Tamako Market Latest	Vader	89	89	100	94	489	1	62	10
	TextBlob	93	93	100	97	517	0	37	8
	TextBlob Plus	93	93	100	97	517	0	37	8
	SentiWordNet	85	85	100	92	461	1	83	17
	SentiWordNet Plus	85	85	100	92	459	2	79	22
	AFIIN	91	91	100	95	504	1	52	5
	AFIIN Plus	91	91	100	95	504	1	52	5
	Flair	80	81	99	89	437	5	106	14
	Flair Plus	73	73	90	80	310	36	115	95
Non Non Biyori Latest	Vader	85	85	99	92	800	8	138	23
	TextBlob	89	89	100	94	867	0	102	0
	TextBlob Plus	89	89	100	94	867	0	102	0
	SentiWo	77	77	100	87	738	2	221	8

Keyword	Library	Accuracy	Precision	Recall	F1	TP	FN	FP	TN
	rdNet								
	SentiWordNet Plus	76	77	100	87	739	1	227	2
	AFIIN	86	86	100	93	838	0	131	0
	AFIIN Plus	86	86	100	93	838	0	131	0
	Flair	75	77	91	83	592	62	180	135
	Flair Plus	75	76	91	83	598	56	189	126
Sewayaki Kitsune no Senkōsan The Helpful Fox Senkōsan Senkōsan Latest	Vader	87	87	100	93	490	1	75	2
	TextBlob	85	86	100	92	485	1	82	0
	TextBlob Plus	85	86	100	92	485	1	82	0
	SentiWordNet	83	84	100	91	474	1	93	0
	SentiWordNet Plus	84	84	100	91	473	1	92	2
	AFIIN	89	89	100	94	498	2	64	4
	AFIIN Plus	89	89	100	94	498	2	63	5
	Flair	68	68	99	81	376	5	174	13
	Flair Plus	71	72	93	81	363	26	141	38
Shoujo Shuumatsu Ryokou Girls' Last Tour Latest	Vader	80	81	98	88	549	11	132	20
	TextBlob	87	88	100	93	619	1	88	4
	TextBlob Plus	87	97	100	93	619	1	89	3
	SentiWordNet	70	72	75	73	300	99	117	196
	SentiWordNet Plus	71	73	77	75	305	92	117	198
	AFIIN	82	82	99	90	571	5	126	10
	AFIIN Plus	81	82	99	90	569	7	126	10
	Flair	63	67	80	73	350	88	173	101
	Flair Plus	65	68	83	75	368	74	195	95
Gakuen Babysitters School Babysitter	Vader	85	85	100	92	425	0	75	0
	TextBlob	91	91	100	95	453	1	44	2
	TextBlob Plus	91	91	100	95	453	1	44	2

Keyword	Library	Accuracy	Precision	Recall	F1	TP	FN	FP	TN
rs Latest	SentiWordNet	78	79	98	87	373	8	102	17
	SentiWordNet Plus	78	79	98	87	369	11	100	20
	AFIIN	87	87	100	93	433	0	67	0
	AFIIN Plus	87	87	100	93	433	0	67	0
	Flair	78	78	100	88	392	0	108	0
	Flair Plus	79	79	100	88	396	0	104	0
Amaama to Inazuma Sweetness & Lightning Latest	Vader	90	89	100	94	514	0	63	29
	TextBlob	92	91	100	95	525	0	51	30
	TextBlob Plus	92	91	100	95	525	0	51	30
	SentiWordNet	81	81	97	88	442	15	101	48
	SentiWordNet Plus	81	82	97	89	443	14	100	49
	AFIIN	88	88	100	94	506	0	70	30
	AFIIN Plus	88	88	100	94	506	0	70	30
	Flair	84	84	100	91	507	0	95	4
	Flair Plus	86	86	100	92	518	1	86	1
Flying Witch Latest	Vader	69	71	74	73	261	92	106	175
	TextBlob	91	91	100	95	576	1	57	0
	TextBlob Plus	91	91	100	95	576	1	57	0
	SentiWordNet	81	82	98	90	507	8	110	9
	SentiWordNet Plus	82	82	98	80	504	9	107	14
	AFIIN	92	92	100	96	580	1	49	4
	AFIIN Plus	91	93	98	95	569	12	44	9
	Flair	82	82	100	90	520	0	112	2
	Flair Plus	84	84	100	91	528	0	104	2
Gochuumon wa Usagi	Vader	94	94	100	97	137	0	8	0
	TextBlob	92	92	100	96	133	0	12	0
	TextBlob	92	92	100	96	133	0	12	0

Keyword	Library	Accuracy	Precision	Recall	F1	TP	FN	FP	TN
desu ka? Gochiusa Top	Plus								
	SentiWordNet	90	90	100	95	130	0	15	0
	SentiWordNet Plus	89	90	98	94	127	2	14	2
	AFIIN	96	96	100	98	138	0	6	1
	AFIIN Plus	96	96	100	98	138	0	6	1
	Flair	83	84	93	89	97	7	18	23
	Flair Plus	76	76	97	85	101	3	32	9
Kami-tachi ni Hiroware ta Otoko By the Grace of the Gods Kamihiro Top	Vader	94	95	99	97	134	2	6	3
	TextBlob	88	89	97	93	117	3	15	10
	TextBlob Plus	88	88	98	93	118	2	16	9
	SentiWordNet	83	87	92	89	107	9	16	13
	SentiWordNet Plus	83	84	97	90	114	3	22	6
	AFIIN	91	92	98	95	127	2	11	5
	AFIIN Plus	90	91	98	95	127	2	12	4
	Flair	88	88	100	94	125	0	17	3
	Flair Plus	83	83	100	91	121	0	24	0
Kimi to Boku. You and Me. Top	Vader	77	74	94	83	74	5	27	31
	TextBlob	89	89	100	94	121	0	15	0
	TextBlob Plus	89	89	100	94	121	0	15	0
	SentiWordNet	83	83	100	91	113	0	23	0
	SentiWordNet Plus	83	83	100	91	113	0	23	0
	AFIIN	79	81	88	84	74	10	18	34
	AFIIN Plus	78	80	86	83	72	12	18	34
	Flair	66	78	73	75	70	26	20	20
	Flair Plus	73	77	85	81	78	14	23	21
Anime Aria Top	Vader	85	87	96	91	122	5	19	18
	TextBlob	90	90	98	94	128	3	14	19

Keyword	Library	Accuracy	Precision	Recall	F1	TP	FN	FP	TN
	TextBlob Plus	89	89	98	93	129	2	16	17
	SentiWordNet	81	80	100	89	129	0	32	3
	SentiWordNet Plus	82	82	100	90	131	0	29	4
	AFIIN	84	85	97	90	122	4	22	16
	AFIIN Plus	84	85	96	90	121	5	22	16
	Flair	62	66	70	67	66	28	35	35
	Flair Plus	74	71	99	83	104	1	42	17
So Ra No Wo To Sound of the Sky Top	Vader	87	79	58	64	17	12	7	106
	TextBlob	94	94	100	97	134	0	8	0
	TextBlob Plus	94	94	100	97	134	0	8	0
	SentiWordNet	93	93	100	96	132	0	10	0
	SentiWordNet Plus	93	93	100	96	132	0	10	0
	AFIIN	85	68	83	73	29	6	16	91
	AFIIN Plus	87	76	77	76	27	8	10	97
	Flair	80	74	62	67	30	18	10	84
	Flair Plus	79	77	71	72	37	15	14	76
Konohan a Kitan Top	Vader	87	87	100	93	120	0	18	2
	TextBlob	91	91	99	95	125	1	12	2
	TextBlob Plus	74	91	81	77	101	23	12	3
	SentiWordNet	81	82	99	90	113	1	25	1
	SentiWordNet Plus	81	82	97	89	108	3	24	5
	AFIIN	86	87	99	93	119	1	18	2
	AFIIN Plus	87	87	100	93	120	0	18	2
	Flair	80	81	97	88	107	3	25	5
	Flair Plus	79	79	100	88	110	0	30	0
Udon no	Vader	60	47	33	38	10	21	11	38

Keyword	Library	Accuracy	Precision	Recall	F1	TP	FN	FP	TN
Kuni no Kinkoro Kemari Poco's Udon World Top	TextBlob	88	88	100	93	70	0	10	0
	TextBlob Plus	88	88	100	93	70	0	10	0
	SentiWordNet	84	84	100	91	66	0	13	1
	SentiWordNet Plus	84	84	100	91	66	0	13	1
	AFIIN	65	65	65	63	26	14	14	26
	AFIIN Plus	62	64	65	63	26	14	16	24
	Flair	79	77	100	87	58	0	17	5
	Flair Plus	82	84	95	89	56	3	11	10
Emiya-san Chi no Kyou no Gohan Today's Menu for the Emiya Family Top	Vader	81	82	89	85	72	9	16	35
	TextBlob	91	91	100	95	120	0	12	0
	TextBlob Plus	91	91	100	95	120	0	12	0
	SentiWordNet	93	93	100	96	123	0	9	0
	SentiWordNet Plus	94	94	100	97	124	0	8	0
	AFIIN	81	79	95	86	79	4	21	28
	AFIIN Plus	81	79	95	86	79	4	21	28
	Flair	73	79	86	82	82	13	23	14
	Flair Plus	77	78	97	87	96	3	27	9
Anime Manaria Friends Mysteria Friends Top	Vader	97	97	100	99	73	0	2	5
	TextBlob	91	92	99	95	73	1	6	0
	TextBlob Plus	91	92	99	95	73	1	6	0
	SentiWordNet	94	93	100	97	70	0	5	5
	SentiWordNet Plus	94	93	100	97	70	0	5	5
	AFIIN	93	92	100	96	69	0	6	5
	AFIIN Plus	91	91	100	5	69	0	7	4
	Flair	84	84	100	91	67	0	13	0
	Flair Plus	84	84	100	91	67	0	13	0

Keyword	Library	Accuracy	Precision	Recall	F1	TP	FN	FP	TN
Anime Super Cub Top	Vader	98	98	100	99	129	0	2	0
	TextBlob	96	96	100	98	126	0	5	0
	TextBlob Plus	96	96	100	98	126	0	5	0
	SentiWordNet	78	79	92	85	83	7	22	19
	SentiWordNet Plus	78	78	94	85	85	5	24	18
	AFINN	98	98	100	99	128	0	3	0
	AFINN Plus	98	98	100	99	128	0	3	0
	Flair	84	84	100	91	110	0	21	0
	Flair Plus	83	84	99	91	108	1	21	1
Danchigai Top	Vader	95	95	100	97	90	0	5	0
	TextBlob	96	96	100	98	90	0	4	1
	TextBlob Plus	95	95	100	97	90	0	5	0
	SentiWordNet	87	87	100	93	83	0	12	0
	SentiWordNet Plus	86	88	98	93	81	2	11	1
	AFINN	96	96	100	98	91	0	4	0
	AFINN Plus	96	96	100	98	91	0	4	0
	Flair	79	83	90	86	63	7	13	12
	Flair Plus	75	75	97	85	67	2	22	4
Anime Ichigo Mashimaro Strawberry Marshmallow Top	Vader	89	88	100	94	83	0	11	2
	TextBlob	90	90	100	95	86	0	10	0
	TextBlob Plus	90	90	100	95	86	0	10	0
	SentiWordNet	83	84	99	91	79	1	15	1
	SentiWordNet Plus	83	84	99	91	79	1	15	1
	AFINN	90	90	100	95	86	0	10	0
	AFINN Plus	90	90	100	95	86	0	10	0
	Flair	67	77	85	74	58	10	22	6
	Flair	71	73	93	82	63	5	23	5

Keyword	Library	Accuracy	Precision	Recall	F1	TP	FN	FP	TN
	Plus								
Hanayata ma Top	Vader	78	79	99	88	74	1	20	0
	TextBlob	91	90	100	95	84	0	9	2
	TextBlob Plus	91	90	100	95	84	0	9	2
	SentiWo rdNet	82	82	100	90	76	0	17	2
	SentiWo rdNet Plus	82	82	100	90	76	0	17	2
	AFIIN	85	85	100	92	81	0	14	0
	AFIIN Plus	85	85	100	92	81	0	14	0
	Flair	79	79	100	88	75	0	20	0
	Flair Plus	76	81	91	86	69	7	16	3
Mitsubos hi Colors Top	Vader	88	88	100	94	52	0	7	0
	TextBlob	97	97	100	98	55	0	2	0
	TextBlob Plus	97	97	100	98	55	0	2	0
	SentiWo rdNet	85	85	100	92	50	0	9	0
	SentiWo rdNet Plus	85	85	100	92	50	0	9	0
	AFIIN	88	88	100	94	52	0	7	0
	AFIIN Plus	88	88	100	94	52	0	7	0
	Flair	87	89	96	92	48	2	6	3
	Flair Plus	90	90	100	95	51	0	6	2
Cool Doji Danshi Play It Cool, Guys Top	Vader	96	96	99	98	158	1	6	0
	TextBlob	96	96	100	98	159	0	6	0
	TextBlob Plus	96	96	100	98	159	0	6	0
	SentiWo rdNet	80	81	98	89	129	2	31	3
	SentiWo rdNet Plus	79	79	100	89	131	0	34	0
	AFIIN	92	93	99	96	151	2	11	1
	AFIIN Plus	92	93	99	96	151	2	11	1

Keyword	Library	Accuracy	Precision	Recall	F1	TP	FN	FP	TN
	Flair	87	87	99	93	139	1	21	4
	Flair Plus	88	88	99	93	142	1	19	3
Shiroku ma Cafe Polar Bear Cafe Top	Vader	88	88	100	93	122	0	17	0
	TextBlob	65	64	73	68	51	19	30	39
	TextBlob Plus	67	65	76	70	53	17	29	40
	SentiWordNet	79	83	94	88	109	7	22	1
	SentiWordNet Plus	75	83	88	85	102	14	21	2
	AFIIN	90	90	100	95	125	0	14	0
	AFIIN Plus	90	90	100	95	125	0	14	0
	Flair	76	76	100	87	106	0	33	0
	Flair Plus	76	76	100	86	105	0	34	0
Amachu Top	Vader	87	87	100	95	90	0	14	0
	TextBlob	87	87	99	93	90	1	13	0
	TextBlob Plus	87	87	99	93	90	1	13	0
	SentiWordNet	83	84	98	91	86	2	16	0
	SentiWordNet Plus	86	87	98	92	86	2	13	3
	AFIIN	89	89	100	94	92	0	11	1
	AFIIN Plus	89	89	100	94	92	0	11	1
	Flair	78	78	100	88	81	0	23	0
	Flair Plus	80	80	100	89	83	0	21	0
Anime Aiura Top	Vader	87	88	99	93	101	1	14	3
	TextBlob	91	91	100	95	106	0	11	2
	TextBlob Plus	92	91	100	96	106	0	10	3
	SentiWordNet	84	85	99	91	99	1	18	1
	SentiWordNet Plus	86	86	99	92	98	1	16	4
	AFIIN	88	88	100	94	102	0	14	3
	AFIIN	88	88	100	94	102	0	14	3

Keyword	Library	Accuracy	Precision	Recall	F1	TP	FN	FP	TN
	Plus								
	Flair	72	76	89	82	76	9	24	10
	Flair Plus	75	75	100	86	89	0	30	0
Yama no Susume Encouragement of Climb Top	Vader	74	69	53	58	24	22	12	75
	TextBlob	91	91	100	95	121	0	12	0
	TextBlob Plus	91	91	100	95	121	0	12	0
	SentiWordNet	71	58	82	68	42	9	30	52
	SentiWordNet Plus	72	59	88	70	45	6	32	50
	AFINN	76	73	79	76	50	13	19	51
	AFINN Plus	79	75	84	79	53	10	18	52
	Flair	84	87	94	90	102	6	16	9
	Flair Plus	81	80	100	89	103	0	25	5
Shounen Maid Shonen Maid Top	Vader	86	86	100	93	113	0	18	0
	TextBlob	90	90	100	95	118	0	13	0
	TextBlob Plus	90	90	100	95	118	0	13	0
	SentiWordNet	86	86	100	92	111	0	18	2
	SentiWordNet Plus	85	85	100	92	111	0	19	1
	AFINN	85	85	99	92	109	1	19	2
	AFINN Plus	85	85	100	92	110	0	19	2
	Flair	78	78	100	87	101	0	29	1
	Flair Plus	78	78	99	87	97	1	28	5
Kyuujitsu no Warumono-san Mr. Villain's Day Off Top	Vader	65	66	93	77	96	7	50	12
	TextBlob	88	88	99	93	139	1	19	6
	TextBlob Plus	88	88	99	93	139	1	19	6
	SentiWordNet	69	64	61	62	43	28	23	71
	SentiWordNet Plus	70	68	56	61	40	31	19	75
	AFINN	82	83	99	90	131	2	27	5

Keyword	Library	Accuracy	Precision	Recall	F1	TP	FN	FP	TN
	AFIIN Plus	82	83	99	90	131	2	27	5
	Flair	70	72	72	72	63	25	25	52
	Flair Plus	67	68	76	72	70	22	33	40
Anime Kemono Friends Top	Vader	91	91	100	95	73	0	7	0
	TextBlob	85	87	97	92	66	2	10	2
	TextBlob Plus	84	87	95	91	65	3	10	2
	SentiWordNet	70	66	80	72	53	13	11	3
	SentiWordNet Plus	82	82	100	90	66	0	14	0
	AFIIN	82	82	100	90	66	0	14	0
	AFIIN Plus	82	82	100	90	66	0	14	0
	Flair	80	79	100	88	60	0	16	4
	Flair Plus	78	77	100	87	61	0	18	1
Honobono Log Top	Vader	91	91	100	95	98	0	10	2
	TextBlob	93	93	100	96	101	0	8	1
	TextBlob Plus	93	93	100	96	101	0	8	1
	SentiWordNet	82	82	99	90	88	1	19	2
	SentiWordNet Plus	83	83	99	90	88	1	18	3
	AFIIN	91	92	99	95	97	1	9	3
	AFIIN Plus	92	92	99	96	97	1	8	4
	Flair	65	67	92	77	66	6	33	5
	Flair Plus	73	75	87	81	62	9	21	18
Hidamari Sketch Top	Vader	87	87	100	93	85	0	13	0
	TextBlob	87	87	100	93	85	0	13	0
	TextBlob Plus	87	87	100	93	85	0	13	0
	SentiWordNet	76	79	93	85	69	5	19	5
	SentiWordNet Plus	76	76	100	86	74	0	24	0

Keyword	Library	Accuracy	Precision	Recall	F1	TP	FN	FP	TN
	AFIIN	85	85	100	92	83	0	15	0
	AFIIN Plus	85	85	100	92	83	0	15	0
	Flair	65	65	98	68	61	1	33	3
	Flair Plus	63	63	100	77	62	0	36	0
Hakumei to Mikochi Hakumei and Mikochi Tiny Little Life in the Woods Top	Vader	80	81	99	89	68	1	16	1
	TextBlob	88	88	100	94	76	0	10	0
	TextBlob Plus	88	88	100	94	76	0	10	0
	SentiWordNet	79	79	100	88	68	0	18	0
	SentiWordNet Plus	79	79	100	88	68	0	18	0
	AFIIN	74	78	94	85	63	4	18	1
	AFIIN Plus	74	78	93	85	62	5	17	2
	Flair	77	77	100	87	66	0	20	0
	Flair Plus	76	76	96	86	65	1	20	0
Hokago Teibo Nisshi, Afterschool Embankment Journal Diary of Our Days at the Breakwater Top	Vader	90	90	100	95	97	0	11	0
	TextBlob	88	90	97	94	94	3	10	1
	TextBlob Plus	88	91	96	93	93	4	9	2
	SentiWordNet	91	91	100	95	96	0	10	2
	SentiWordNet Plus	91	91	99	95	95	1	9	3
	AFIIN	88	88	100	94	95	0	13	0
	AFIIN Plus	87	88	99	93	94	1	13	0
	Flair	80	82	94	88	78	5	17	8
	Flair Plus	84	84	100	91	86	0	17	5
Anime Slow Loop Top	Vader	91	91	100	95	97	0	10	0
	TextBlob	79	74	78	76	35	10	13	49
	TextBlob Plus	78	72	78	75	35	10	14	48
	SentiWordNet	88	88	100	94	94	0	13	0
	SentiWordNet	87	87	100	93	93	0	14	0

Keyword	Library	Accuracy	Precision	Recall	F1	TP	FN	FP	TN
	Plus								
	AFIIN	93	92	100	96	98	0	8	1
	AFIIN Plus	93	92	100	96	98	0	8	1
	Flair	78	72	90	80	45	5	18	39
	Flair Plus	74	66	94	77	47	3	25	32
Kamichu Top	Vader	83	83	100	91	102	0	21	0
	TextBlob	86	86	100	93	106	0	17	0
	TextBlob Plus	86	86	100	93	106	0	17	0
	SentiWordNet	80	81	97	88	95	3	22	3
	SentiWordNet Plus	80	80	100	89	98	0	25	0
	AFIIN	82	82	100	90	101	0	22	0
	AFIIN Plus	82	82	100	90	101	0	22	0
	Flair	73	73	100	95	90	0	33	0
	Flair Plus	77	79	94	86	85	5	23	10
Ikoku Meiro no Croisée The Animation Croisée in a Foreign Labyrinth The Animation Top	Vader	67	66	57	60	24	18	12	36
	TextBlob	84	84	100	91	73	0	14	3
	TextBlob Plus	84	84	100	91	73	0	14	3
	SentiWordNet	78	77	100	87	67	0	20	3
	SentiWordNet Plus	78	77	100	87	67	0	20	3
	AFIIN	67	67	61	63	26	17	13	34
	AFIIN Plus	76	77	70	72	30	13	9	38
	Flair	59	60	63	60	29	17	20	24
	Flair Plus	59	62	66	63	32	17	20	21
Atarashii Joushi wa Do Tennen My New Boss Is Goofy	Vader	87	87	100	93	125	0	19	0
	TextBlob	97	97	100	98	140	0	4	0
	TextBlob Plus	97	97	100	98	140	0	4	0
	SentiWordNet	90	89	100	94	126	0	15	3
	SentiWo	89	89	99	94	125	1	15	3

Keyword	Library	Accuracy	Precision	Recall	F1	TP	FN	FP	TN
Top	rdNet Plus								
	AFIIN	73	64	39	48	18	28	11	87
	AFIIN Plus	69	53	44	47	20	26	19	79
	Flair	77	77	100	87	110	0	33	1
	Flair Plus	82	83	98	90	114	2	24	4
Anime Tamyura Top	Vader	87	86	99	92	137	1	23	17
	TextBlob	93	93	100	96	162	0	13	3
	TextBlob Plus	93	93	100	96	162	0	13	3
	SentiWordNet	85	87	98	92	149	3	23	3
	SentiWordNet Plus	86	87	98	92	149	3	22	4
	AFIIN	89	89	98	93	132	3	16	27
	AFIIN Plus	88	87	100	93	135	0	21	22
	Flair	80	79	97	87	120	4	32	22
	Flair Plus	75	75	100	86	134	0	44	0
Yakumo Yaku nara Mug Cup mo Let's Make a Mug Too Top	Vader	91	91	100	96	85	0	8	0
	TextBlob	92	94	99	96	86	1	6	0
	TextBlob Plus	91	94	97	95	84	3	5	1
	SentiWordNet	74	95	35	48	13	23	1	56
	SentiWordNet Plus	74	95	35	48	13	23	1	56
	AFIIN	95	95	100	97	88	0	5	0
	AFIIN Plus	95	95	100	97	88	0	5	0
	Flair	64	74	57	61	28	21	12	32
	Flair Plus	72	70	95	80	52	3	23	15
Anime Yokohama Shopping Log, Yokohama	Vader	86	87	99	92	86	1	13	1
	TextBlob	87	89	98	93	87	2	11	1
	TextBlob Plus	87	89	98	93	87	2	11	1
	SentiWordNet	79	81	96	88	76	3	18	4

Keyword	Library	Accuracy	Precision	Recall	F1	TP	FN	FP	TN
ma Shopping Trip Top	SentiWordNet Plus	79	81	96	88	75	3	18	5
	AFIIN	88	88	100	94	89	0	12	0
	AFIIN Plus	88	88	100	94	89	0	12	0
	Flair	63	67	85	75	56	10	27	8
	Flair Plus	69	70	99	82	69	1	30	1
NieA Under 7 NieA_7 Top	Vader	85	85	100	92	104	0	19	5
	TextBlob	90	90	100	94	110	0	13	5
	TextBlob Plus	90	90	100	94	110	0	13	5
	SentiWordNet	84	84	100	91	103	0	20	5
	SentiWordNet Plus	84	84	99	91	102	1	20	5
	AFIIN	88	88	100	94	108	0	15	5
	AFIIN Plus	88	88	100	94	108	0	15	5
	Flair	71	71	98	82	85	2	35	6
	Flair Plus	70	69	100	82	87	0	39	2
Rokuhou-dou Yotsuiro Biyori Yotsuiro Biyori Top	Vader	98	98	100	99	84	0	2	0
	TextBlob	95	95	100	98	82	0	4	0
	TextBlob Plus	95	95	100	98	82	0	4	0
	SentiWordNet	92	92	100	96	79	0	7	0
	SentiWordNet Plus	92	92	100	96	79	0	7	0
	AFIIN	97	97	100	98	83	0	3	0
	AFIIN Plus	97	97	100	98	83	0	3	0
	Flair	92	92	100	96	79	0	7	0
	Flair Plus	94	94	100	97	69	0	5	12
Anime Sketchbook full color's Top	Vader	98	98	100	99	84	0	2	0
	TextBlob	95	95	100	98	82	0	4	0
	TextBlob Plus	95	95	100	98	82	0	4	0
	SentiWordNet	92	92	100	96	79	0	7	0

Keyword	Library	Accuracy	Precision	Recall	F1	TP	FN	FP	TN
	rdNet								
	SentiWordNet Plus	92	92	100	96	79	0	7	0
	AFIIN	97	97	100	98	83	0	3	0
	AFIIN Plus	97	97	100	98	83	0	3	0
	Flair	92	92	100	96	79	0	7	0
	Flair Plus	94	94	100	97	69	0	5	12
Moomin, Tanoshii Muumin Ikka Top	Vader	88	88	100	94	80	0	11	0
	TextBlob	95	95	100	97	86	0	5	0
	TextBlob Plus	95	95	100	97	86	0	5	0
	SentiWordNet	79	80	99	88	72	1	18	0
	SentiWordNet Plus	80	80	100	89	73	0	18	0
	AFIIN	92	92	100	96	84	0	7	0
	AFIIN Plus	92	92	100	96	84	0	7	0
	Flair	80	81	99	89	70	1	17	3
	Flair Plus	77	77	100	87	70	0	21	0
Mahoutsukai ni Taisetsu na Koto Someday's Dreamers Top	Vader	88	88	100	94	80	0	11	0
	TextBlob	95	95	100	97	86	0	5	0
	TextBlob Plus	95	95	100	97	86	0	5	0
	SentiWordNet	79	80	99	88	72	1	18	0
	SentiWordNet Plus	80	80	100	89	73	0	18	0
	AFIIN	92	92	100	96	84	0	7	0
	AFIIN Plus	92	92	100	96	84	0	7	0
	Flair	80	81	99	89	70	1	17	3
	Flair Plus	77	77	100	87	70	0	21	0
Mahoraba: Heartful Days	Vader	92	91	100	95	104	0	10	6
	TextBlob	93	94	98	96	102	2	7	9
	TextBlob Plus	91	91	100	95	104	0	11	5

Keyword	Library	Accuracy	Precision	Recall	F1	TP	FN	FP	TN
Top	SentiWordNet	83	87	94	91	96	6	14	4
	SentiWordNet Plus	85	88	95	92	97	5	13	5
	AFIIN	89	90	98	94	103	2	11	4
	AFIIN Plus	90	91	99	95	104	1	11	4
	Flair	91	93	97	95	105	3	8	4
	Flair Plus	93	93	100	96	108	0	8	4
Fuujin Monogatari Windy Tales Top	Vader	90	91	99	95	89	1	9	0
	TextBlob	90	90	100	95	89	0	10	0
	TextBlob Plus	90	90	100	95	89	0	10	0
	SentiWordNet	83	84	98	90	81	2	15	1
	SentiWordNet Plus	84	84	100	91	83	0	16	0
	AFIIN	91	91	100	95	90	0	9	0
	AFIIN Plus	90	91	99	95	104	1	11	4
	Flair	83	83	99	90	80	1	16	2
Maiko-san Chino Mekanaisan The caterer at the Maiko Manor Top	Flair Plus	85	85	100	92	82	0	15	2
	Vader	84	87	87	87	88	13	13	48
	TextBlob	96	96	100	98	154	0	7	1
	TextBlob Plus	96	96	100	98	154	0	6	2
	SentiWordNet	92	92	100	96	149	0	13	0
	SentiWordNet Plus	92	92	100	96	149	0	13	0
	AFIIN	87	91	89	90	93	11	10	48
	AFIIN Plus	86	90	88	89	92	12	10	48
	Flair	88	88	99	93	135	1	18	8
Anime One Off Top	Flair Plus	86	86	100	93	136	0	22	4
	Vader	76	75	100	86	64	0	21	1
	TextBlob	77	77	100	87	65	0	20	1
	TextBlob	77	77	100	87	65	0	20	1

Keyword	Library	Accuracy	Precision	Recall	F1	TP	FN	FP	TN
	Plus								
	SentiWordNet	72	73	98	84	61	1	23	1
	SentiWordNet Plus	69	71	95	81	58	3	24	1
	AFIIN	76	76	100	86	65	0	21	0
	AFIIN Plus	76	76	100	86	65	0	21	0
	Flair	72	0	0	0	0	24	0	62
	Flair Plus	83	16	12	13	3	25	7	51
Anime Mokke Top	Vader	85	85	100	92	47	0	8	0
	TextBlob	85	85	100	92	47	0	8	0
	TextBlob Plus	85	85	100	92	47	0	8	0
	SentiWordNet	75	75	100	85	41	0	14	0
	SentiWordNet Plus	75	75	100	85	41	0	14	0
	AFIIN	91	91	100	95	50	0	5	0
	AFIIN Plus	91	91	100	95	50	0	5	0
	Flair	55	56	56	56	16	13	12	14
	Flair Plus	55	52	68	57	19	8	7	11
Rilakku ma to Kaorusan Rilakku ma and Kaoru Top	Vader	87	88	99	93	146	2	21	10
	TextBlob	91	92	99	95	157	2	14	6
	TextBlob Plus	91	92	99	95	157	2	14	6
	SentiWordNet	86	88	97	92	152	4	21	2
	SentiWordNet Plus	87	88	99	93	155	1	22	1
	AFIIN	91	91	99	95	149	1	15	14
	AFIIN Plus	90	90	99	95	149	1	16	13
	Flair	81	80	100	89	138	0	34	7
	Flair Plus	81	81	99	89	136	1	33	9
Binchoutan	Vader	87	87	100	93	110	0	16	0
	TextBlob	87	87	100	93	109	0	17	0

Keyword	Library	Accuracy	Precision	Recall	F1	TP	FN	FP	TN
Binchotan Top	TextBlob Plus	87	87	100	93	109	0	17	0
	SentiWordNet	79	82	94	88	97	6	21	2
	SentiWordNet Plus	84	85	89	91	102	2	18	4
	AFIIN	86	87	98	92	108	2	16	0
	AFIIN Plus	87	87	99	93	109	1	16	0
	Flair	71	77	84	80	74	14	23	15
	Flair Plus	72	73	95	83	85	4	31	6
Yoru no Kuni, Night World, Night Country	Vader	93	93	100	97	114	0	8	0
	TextBlob	98	98	100	99	119	0	3	0
	TextBlob Plus	98	98	100	99	119	0	3	0
	SentiWordNet	93	93	100	96	113	0	9	0
	SentiWordNet Plus	93	93	100	96	113	0	9	0
	AFIIN	96	96	100	98	117	0	5	0
	AFIIN Plus	96	96	100	98	117	0	5	0
	Flair	88	88	100	93	106	0	15	1
	Flair Plus	90	90	100	95	110	0	12	0
Kokoro Toshokan Kokoro Library Top	Vader	87	87	100	93	48	0	7	0
	TextBlob	93	93	100	96	51	0	4	0
	TextBlob Plus	93	93	100	96	51	0	4	0
	SentiWordNet	85	85	100	92	47	0	8	0
	SentiWordNet Plus	85	85	100	92	47	0	8	0
	AFIIN	89	89	100	94	49	0	6	0
	AFIIN Plus	89	89	100	94	49	0	6	0
	Flair	69	69	100	82	38	0	17	0
	Flair Plus	71	71	100	83	39	0	16	0
Kinoko	Vader	81	80	99	89	122	1	30	9

Keyword	Library	Accuracy	Precision	Recall	F1	TP	FN	FP	TN
Inu Kinokoin u Mushroom Pup Top	TextBlob	91	91	100	95	148	0	14	0
	TextBlob Plus	91	91	100	95	147	0	15	0
	SentiWordNet	85	86	97	91	122	4	20	16
	SentiWordNet Plus	86	87	96	91	121	5	18	18
	AFIIN	89	89	100	94	140	0	18	4
	AFIIN Plus	88	89	99	93	137	3	17	5
	Flair	75	75	97	84	108	3	37	14
	Flair Plus	75	76	96	85	112	5	36	9

3. Hasil Pengambilan 5 Data Pertama Anime menggunakan API Anilist

Data	Nilai
Romaji Title	Shingeki no Kyojin
English Title	Attack on Titan
Anilist link	https://anilist.co/anime/16498/Attack-on-Titan/
Genres	Action, Drama, Fantasy, Mystery
Tags	Revenge 97%, Kaiju 93%, Male Protagonist 89%, Military 88%, Tragedy 88%, Post-Apocalyptic 84%, Primarily Teen Cast 84%, Super Power 82%, Steampunk 79%, Henshin 79%, Swordplay 76%, Gore 75%, Memory Manipulation 75%, Dystopian 74%, Ensemble Cast 73%, Orphan 73%, Kuudere 70%, Espionage 68%, Cannibalism 66%, Coming of Age 62%, Shounen 61%, Survival 60%, Medieval 60%, Vore 60%, Time Skip 59%, Amnesia 58%, Rural 51%, Suicide 50%, CGI 46%, Primarily Male Cast 44%, Adoption 35%
Rating	84
Original Work Finished	Yes
Fiction	Yes
LGBT	No
Extreme Violence	Yes
Requires Thinking	Yes
Streaming Links	Crunchyroll http://www.crunchyroll.com/attack-on-titan , Hulu http://www.hulu.com/attack-on-titan , Crunchyroll https://www.crunchyroll.com/attack-on-titan-dubs

Data	Nilai
Romaji Title	Kimetsu no Yaiba
English Title	Demon Slayer: Kimetsu no Yaiba
Anilist link	https://anilist.co/anime/101922/Demon-Slayer-Kimetsu-no-Yaiba/
Genres	Action, Adventure, Drama, Fantasy, Supernatural
Tags	Demons 96%, Shounen 92%, Swordplay 88%, Male Protagonist 81%, Vampire 79%, Tragedy 78%, Monster Girl 75%, Primarily Teen Cast 75%, Revenge 73%, Travel 73%, Orphan 68%, CGI 67%, Gore 66%, Primarily Male Cast 65%, Mythology 65%, Rural 60%, Slapstick 60%, Historical 58%, Rotoscoping 56%, Super Power 56%, Body Horror 55%, Curses 50%, Chibi 49%, Time Skip 42%, Snowscape 35%, Animals 20%, Food 20%
Rating	82
Original Work Finished	Yes
Fiction	Yes
LGBT	No
Extreme Violence	Yes
Requires Thinking	Yes
Streaming Links	Crunchyroll https://www.crunchyroll.com/demon-slayer-kimetsu-no-yaiba , Hulu https://www.hulu.com/demon-slayer-kimetsu-no-yaiba , Netflix https://www.netflix.com/title/81091393

Data	Nilai
Romaji Title	DEATH NOTE
English Title	Death Note
Anilist link	https://anilist.co/anime/1535/Death-Note/
Genres	Mystery, Psychological, Supernatural, Thriller
Tags	Crime 96%, Detective 95%, Anti-Hero 94%, Male Protagonist 89%, Fugitive 88%, Police 87%, Philosophy 84%, Primarily Adult Cast 80%, Kuudere 79%, Gods 77%, Memory Manipulation 77%, Urban Fantasy 75%, Primarily Male Cast 71%, Urban 71%, Unrequited Love 71%, Amnesia 65%, Yandere 65%, Espionage 64%, Shounen 64%, Suicide 54%, Tennis 53%, Acting 47%, Time Skip 46%, Assassins 42%, Achronological Order 20%, Asexual 10%
Rating	84
Original Work Finished	Yes
Fiction	Yes
LGBT	No

Data	Nilai
Extreme Violence	Yes
Requires Thinking	Yes
Streaming Links	Hulu http://www.hulu.com/death-note , Netflix https://www.netflix.com/title/70204970 , YouTube https://youtube.com/playlist?list=PLDDFkfLheQ9hziWtnCSla1DicLZkllriY

Data	Nilai
Romaji Title	Jujutsu Kaisen
English Title	JUJUTSU KAISEN
Anilist link	https://anilist.co/anime/113415/JUJUTSU-KAISEN/
Genres	Action, Drama, Supernatural
Tags	Urban Fantasy 93%, Shounen 92%, Youkai 89%, Super Power 88%, Curses 86%, Male Protagonist 85%, Primarily Teen Cast 84%, Exorcism 84%, Dissociative Identities 80%, Demons 80%, Orphan 80%, Urban 79%, Body Horror 75%, Magic 72%, Martial Arts 71%, Ensemble Cast 69%, Primarily Male Cast 66%, Gore 65%, Shapeshifting 64%, Swordplay 61%, School 60%, Anthropomorphism 57%, Rotoscoping 56%, Slapstick 50%, Mythology 50%, Surreal Comedy 40%, Twins 40%, Boarding School 40%, Baseball 28%
Rating	85
Original Work Finished	Yes
Fiction	Yes
LGBT	No
Extreme Violence	Yes
Requires Thinking	Yes
Streaming Links	Crunchyroll https://www.crunchyroll.com/jujutsu-kaisen , Netflix https://www.netflix.com/title/81278456 , Bilibili TV https://www.bilibili.tv/media/37738 , YouTube https://youtube.com/playlist?list=PLxSscENep7JisDU6GAJuyNpVwDvCm-f3J , Hulu https://www.hulu.com/series/jujutsu-kaisen-382ec8bf-3650-4cde-94db-ecd18665f9e0

Data	Nilai
Romaji Title	Boku no Hero Academia
English Title	My Hero Academia

Data	Nilai
Anilist link	https://anilist.co/anime/21459/My-Hero-Academia/
Genres	Action, Adventure, Comedy
Tags	Super Power 98%, Superhero 97%, Shounen 90%, Primarily Teen Cast 89%, Male Protagonist 86%, Urban Fantasy 82%, School 81%, Ensemble Cast 75%, Coming of Age 71%, Bullying 66%, Cultivation 65%, Primarily Male Cast 64%, Prison 35%, Yuri 10%
Rating	76
Original Work Finished	No
Fiction	Yes
LGBT	Yes
Extreme Violence	Yes
Requires Thinking	No
Streaming Links	Hulu http://www.hulu.com/my-hero-academia , Crunchyroll http://www.crunchyroll.com/my-hero-academia , Netflix https://www.netflix.com/title/80135674

4. Hasil akurasi KNN dan Naive Bayes jika anime dipisah <60 jelek, 60-80 bagus, >=80 sangat bagus

KNN Recommendations for: shingeki no kyojin

KNN - Mencari anime termirip dengan: Shingeki no Kyojin (kecocokan: 100%)

	Recommended Anime	Similarity	Rating
0	Shingeki no Kyojin Season 2	27.88%	84.0
1	Shingeki no Kyojin: Chronicle	25.24%	76.0
2	Shingeki no Kyojin Season 3 Part 2	24.94%	89.0
3	Shingeki no Kyojin OVA	24.11%	76.0
4	Shingeki no Kyojin Zenpen: Guren no Yumiya	23.63%	75.0

Top Genres

Top

Tags

0 Action, Drama, Fantasy Survival, Post-Apocalyptic, Shapeshifting

1 Action, Drama, Fantasy Kaiju, Post-Apocalyptic, Tragedy

2	Action, Drama, Fantasy Conspiracy	Dystopian, Military,
3	Action, Drama, Fantasy Power	Kaiju, Military, Super
4	Action, Drama, Fantasy Kaiju	Super Power, Medieval,

Naive Bayes Recommendations for: shingeki no kyojin

Naive Bayes - Mencari anime termirip dengan: Shingeki no Kyojin
(kecocokan: 100%)

	Recommended Anime	Similarity	Rating \
0	Shingeki no Kyojin Season 2	100.00%	84.0
1	Shingeki no Kyojin Season 3 Part 2	93.61%	89.0
2	Shingeki no Kyojin: Chronicle	91.64%	76.0
3	Shingeki no Kyojin Season 3	88.98%	85.0
4	Shingeki no Kyojin: The Final Season	88.95%	86.0

	Top Genres	Top
Tags		
0	Action, Drama, Fantasy Survival, Post-Apocalyptic, Shapeshifting	
1	Action, Drama, Fantasy Dystopian, Military, Conspiracy	
2	Action, Drama, Fantasy Kaiju, Post-Apocalyptic, Tragedy	
3	Action, Drama, Fantasy Conspiracy, Military, Tragedy	
4	Action, Drama, Fantasy War, Tragedy, Military	

Evaluating KNN...

Best k for KNN: 4 (Average Accuracy: 78.19%)

KNN Fold 1 Accuracy: 77.89%

KNN Fold 1 Confusion Matrix:

	Bad	Good	Very Good
Bad	1191	128	3
Good	309	463	5

Very Good 1 26 9

KNN Fold 2 Accuracy: 77.42%

KNN Fold 2 Confusion Matrix:

	Bad	Good	Very Good
Bad	1199	149	1
Good	296	450	7
Very Good	6	23	4

KNN Fold 3 Accuracy: 79.63%

KNN Fold 3 Confusion Matrix:

	Bad	Good	Very Good
Bad	1217	146	0
Good	252	475	3
Very Good	7	27	8

KNN Fold 4 Accuracy: 78.31%

KNN Fold 4 Confusion Matrix:

	Bad	Good	Very Good
Bad	1186	159	2
Good	277	475	2
Very Good	5	18	11

KNN Fold 5 Accuracy: 77.28%

KNN Fold 5 Confusion Matrix:

	Bad	Good	Very Good
Bad	1183	137	0
Good	302	460	6
Very Good	7	33	7

KNN Fold 6 Accuracy: 79.57%

KNN Fold 6 Confusion Matrix:

	Bad	Good	Very Good
Bad	1214	153	2
Good	240	476	5
Very Good	5	31	8

KNN Fold 7 Accuracy: 75.91%

KNN Fold 7 Confusion Matrix:

	Bad	Good	Very Good
Bad	1177	185	0
Good	297	431	6
Very Good	4	22	12

KNN Fold 8 Accuracy: 77.93%

KNN Fold 8 Confusion Matrix:

	Bad	Good	Very Good
Bad	1198	147	1
Good	293	453	4
Very Good	7	19	12

KNN Fold 9 Accuracy: 78.07%

KNN Fold 9 Confusion Matrix:

	Bad	Good	Very Good
Bad	1215	148	1
Good	283	438	4
Very Good	4	28	13

KNN Fold 10 Accuracy: 79.85%

KNN Fold 10 Confusion Matrix:

	Bad	Good	Very Good
Bad	1235	141	0
Good	255	463	2
Very Good	4	28	6

KNN Confusion Matrix:

	Bad	Good	Very Good
Bad	12015	1493	10
Good	2804	4584	44
Very Good	50	255	90

KNN Sample Predictions (Title, Actual, Predicted, Correct):

Title: Tenshi no Pants, Actual: Bad, Predicted: Bad, Correct: True

Title: Ponkotsuland Saga, Actual: Bad, Predicted: Bad, Correct: True

Title: Little El Cid no Bouken, Actual: Bad, Predicted: Bad, Correct: True

Title: Marine Snow no Densetsu, Actual: Bad, Predicted: Bad, Correct: True

Title: Golden Ray, Actual: Good, Predicted: Good, Correct: True

Title: GOSICK, Actual: Good, Predicted: Good, Correct: True

Title: Kafun Shoujo Chuuihou! THE ANIMATION, Actual: Good, Predicted: Bad, Correct: False

Title: Hokuto no Ken, Actual: Good, Predicted: Good, Correct: True

Title: Cross Game, Actual: Very Good, Predicted: Good, Correct: False

Title: HELLSING OVA, Actual: Very Good, Predicted: Good, Correct: False

KNN Evaluation Metrics (10-Fold CV):

Accuracy: 78.19%

Precision: 0.78

Recall: 0.78

F1-Score: 0.77

Average Similarity: 0.5802

Coverage: 0.8955

Diversity: 2.8457

Evaluating Naive Bayes...

NAIVE_BAYES Fold 1 Accuracy: 76.11%

NAIVE_BAYES Fold 1 Confusion Matrix:

	Bad	Good	Very Good
Bad	1177	137	8
Good	319	430	28
Very Good	2	16	18

NAIVE_BAYES Fold 2 Accuracy: 76.96%

NAIVE_BAYES Fold 2 Confusion Matrix:

	Bad	Good	Very Good
Bad	1222	124	3
Good	310	405	38
Very Good	2	15	16

NAIVE_BAYES Fold 3 Accuracy: 77.52%

NAIVE_BAYES Fold 3 Confusion Matrix:

	Bad	Good	Very Good
Bad	1211	149	3
Good	273	420	37
Very Good	1	17	24

NAIVE_BAYES Fold 4 Accuracy: 76.49%

NAIVE_BAYES Fold 4 Confusion Matrix:

	Bad	Good	Very Good
Bad	1201	142	4
Good	306	413	35
Very Good	1	14	19

NAIVE_BAYES Fold 5 Accuracy: 77.42%

NAIVE_BAYES Fold 5 Confusion Matrix:

	Bad	Good	Very Good
Bad	1192	123	5
Good	301	434	33
Very Good	1	19	27

NAIVE_BAYES Fold 6 Accuracy: 76.29%

NAIVE_BAYES Fold 6 Confusion Matrix:

	Bad	Good	Very Good
Bad	1209	153	7
Good	289	400	32
Very Good	2	23	19

NAIVE_BAYES Fold 7 Accuracy: 74.51%

NAIVE_BAYES Fold 7 Confusion Matrix:

	Bad	Good	Very Good
Bad	1176	179	7
Good	311	394	29
Very Good	1	17	20

NAIVE_BAYES Fold 8 Accuracy: 77.27%

NAIVE_BAYES Fold 8 Confusion Matrix:

	Bad	Good	Very Good
Bad	1211	128	7
Good	302	417	31
Very Good	0	17	21

NAIVE_BAYES Fold 9 Accuracy: 77.23%

NAIVE_BAYES Fold 9 Confusion Matrix:

	Bad	Good	Very Good
Bad	1214	142	8
Good	269	412	44
Very Good	0	23	22

NAIVE_BAYES Fold 10 Accuracy: 77.32%

NAIVE_BAYES Fold 10 Confusion Matrix:

	Bad	Good	Very Good
Bad	1215	157	4
Good	265	417	38
Very Good	2	18	18

NAIVE_BAYES Confusion Matrix:

	Bad	Good	Very Good
Bad	12028	1434	56
Good	2945	4142	345
Very Good	12	179	204

NAIVE_BAYES Sample Predictions (Title, Actual, Predicted, Correct):

Title: Tenshi no Pants, Actual: Bad, Predicted: Bad, Correct: True

Title: Ponkotsuland Saga, Actual: Bad, Predicted: Bad, Correct: True

Title: Little El Cid no Bouken, Actual: Bad, Predicted: Bad, Correct: True

Title: Marine Snow no Densetsu, Actual: Bad, Predicted: Bad, Correct: True

Title: Golden Ray, Actual: Good, Predicted: Bad, Correct: False

Title: GOSICK, Actual: Good, Predicted: Very Good, Correct: False

Title: Kafun Shoujo Chuuihou! THE ANIMATION, Actual: Good, Predicted: Bad, Correct: False

Title: Hokuto no Ken, Actual: Good, Predicted: Good, Correct: True

Title: Cross Game, Actual: Very Good, Predicted: Good, Correct: False

Title: HELLSING OVA, Actual: Very Good, Predicted: Very Good, Correct: True

NAIVE_BAYES Evaluation Metrics (10-Fold CV):

Accuracy: 76.71%

Precision: 0.77

Recall: 0.77

F1-Score: 0.76

Average Similarity: 0.3510

Coverage: 1.0000

Diversity: 4.0775

5. Hasil akurasi KNN dan Naive Bayes jika anime dipisah <65 jelek, 65-80 bagus, >=80 sangat bagus

KNN Recommendations for: shingeki no kyojin

KNN - Mencari anime termirip dengan: Shingeki no Kyojin (kecocokan: 100%)

	Recommended Anime	Similarity	Rating
\			
0	Shingeki no Kyojin Season 2	27.88%	84.0
1	Shingeki no Kyojin: Chronicle	25.24%	76.0
2	Shingeki no Kyojin Season 3 Part 2	24.94%	89.0
3	Shingeki no Kyojin OVA	24.11%	76.0
4	Shingeki no Kyojin Zenpen: Guren no Yumiya	23.63%	75.0

	Top Genres	Top
Tags		
0	Action, Drama, Fantasy	Survival, Post-Apocalyptic, Shapeshifting
1	Action, Drama, Fantasy	Kaiju, Post-Apocalyptic, Tragedy
2	Action, Drama, Fantasy	Dystopian, Military, Conspiracy
3	Action, Drama, Fantasy	Kaiju, Military, Super Power
4	Action, Drama, Fantasy	Super Power, Medieval, Kaiju

Naive Bayes Recommendations for: shingeki no kyojin

Naive Bayes - Mencari anime termirip dengan: Shingeki no Kyojin
(kecocokan: 100%)

	Recommended Anime	Similarity	Rating \
0	Shingeki no Kyojin Season 2	100.00%	84.0
1	Shingeki no Kyojin Season 3 Part 2	93.22%	89.0
2	Shingeki no Kyojin: Chronicle	89.88%	76.0
3	Shingeki no Kyojin: The Final Season	88.99%	86.0
4	Shingeki no Kyojin Season 3	88.59%	85.0

	Top Genres	Top Tags
0	Action, Drama, Fantasy	Survival, Post-Apocalyptic, Shapeshifting
1	Action, Drama, Fantasy	Dystopian, Military, Conspiracy
2	Action, Drama, Fantasy	Kaiju, Post-Apocalyptic, Tragedy
3	Action, Drama, Fantasy	War, Tragedy, Military
4	Action, Drama, Fantasy	Conspiracy, Military, Tragedy

Evaluating KNN...

Best k for KNN: 4 (Average Accuracy: 82.23%)

KNN Fold 1 Accuracy: 82.11%

KNN Fold 1 Confusion Matrix:

	Bad	Good	Very Good
Bad	1538	84	3
Good	263	206	5
Very Good	2	25	9

KNN Fold 2 Accuracy: 80.42%

KNN Fold 2 Confusion Matrix:

	Bad	Good	Very Good
Bad	1530	85	1
Good	296	183	7
Very Good	8	21	4

KNN Fold 3 Accuracy: 82.81%

KNN Fold 3 Confusion Matrix:

	Bad	Good	Very Good
Bad	1529	88	0
Good	242	231	3
Very Good	13	21	8

KNN Fold 4 Accuracy: 84.45%

KNN Fold 4 Confusion Matrix:

	Bad	Good	Very Good
Bad	1553	83	2
Good	222	239	2
Very Good	6	17	11

KNN Fold 5 Accuracy: 81.36%

KNN Fold 5 Confusion Matrix:

	Bad	Good	Very Good
Bad	1520	83	1
Good	269	210	5
Very Good	9	31	7

KNN Fold 6 Accuracy: 81.16%

KNN Fold 6 Confusion Matrix:

	Bad	Good	Very Good
Bad	1526	100	2
Good	259	198	5
Very Good	9	27	8

KNN Fold 7 Accuracy: 81.54%

KNN Fold 7 Confusion Matrix:

	Bad	Good	Very Good
Bad	1535	104	1
Good	258	193	5
Very Good	7	19	12

KNN Fold 8 Accuracy: 82.90%

KNN Fold 8 Confusion Matrix:

	Bad	Good	Very Good
--	-----	------	-----------

Bad	1528	77	1
Good	257	229	4
Very Good	12	14	12

KNN Fold 9 Accuracy: 81.87%

KNN Fold 9 Confusion Matrix:

	Bad	Good	Very Good
Bad	1522	96	1
Good	254	212	4
Very Good	7	25	13

KNN Fold 10 Accuracy: 83.65%

KNN Fold 10 Confusion Matrix:

	Bad	Good	Very Good
Bad	1543	92	0
Good	223	236	2
Very Good	8	24	6

KNN Confusion Matrix:

	Bad	Good	Very Good
Bad	15324	892	12
Good	2543	2137	42
Very Good	81	224	90

KNN Sample Predictions (Title, Actual, Predicted, Correct):

Title: Four Seasons: Fuuka to Nanami, Actual: Bad, Predicted: Bad, Correct: True

Title: Meitantei Conan: Keisatsu Gakkou-hen - Wild Police Story CASE. Furuya Rei, Actual: Bad, Predicted: Good, Correct: False

Title: Digimon Frontier: Ornismon Fukkatsu!!, Actual: Bad, Predicted: Good, Correct: False

Title: Pocket Monsters Best Wishes!: Victini to Shiroki Eiyuu Reshiram, Actual: Bad, Predicted: Bad, Correct: True

Title: Sabagebu! OVA SPECIAL MISSION, Actual: Good, Predicted: Good, Correct: True

Title: Chihayafuru 2: Waga Mi Yo ni Furu Nagame Seshi Ma ni, Actual: Good, Predicted: Bad, Correct: False

Title: No Guns Life 2, Actual: Good, Predicted: Bad, Correct: False

Title: Cuticle Tantei Inaba, Actual: Good, Predicted: Good,
Correct: True
Title: Sasaki to Miyano, Actual: Very Good, Predicted: Good,
Correct: False
Title: Made in Abyss: Fukaki Tamashii no Reimei, Actual: Very Good,
Predicted: Bad, Correct: False

KNN Evaluation Metrics (10-Fold CV):

Accuracy: 82.23%

Precision: 0.81

Recall: 0.82

F1-Score: 0.81

Average Similarity: 0.5802

Coverage: 0.8955

Diversity: 2.8457

Evaluating Naive Bayes...

NAIVE_BAYES Fold 1 Accuracy: 81.08%

NAIVE_BAYES Fold 1 Confusion Matrix:

	Bad	Good	Very Good
Bad	1466	150	9
Good	200	252	22
Very Good	3	20	13

NAIVE_BAYES Fold 2 Accuracy: 82.58%

NAIVE_BAYES Fold 2 Confusion Matrix:

	Bad	Good	Very Good
Bad	1495	115	6
Good	204	253	29
Very Good	5	13	15

NAIVE_BAYES Fold 3 Accuracy: 82.30%

NAIVE_BAYES Fold 3 Confusion Matrix:

	Bad	Good	Very Good
Bad	1465	146	6
Good	179	270	27
Very Good	1	19	22

NAIVE_BAYES Fold 4 Accuracy: 83.14%

NAIVE_BAYES Fold 4 Confusion Matrix:

	Bad	Good	Very Good
Bad	1508	126	4
Good	185	251	27
Very Good	3	15	16

NAIVE_BAYES Fold 5 Accuracy: 82.81%

NAIVE_BAYES Fold 5 Confusion Matrix:

	Bad	Good	Very Good
Bad	1480	120	4
Good	194	261	29
Very Good	2	18	27

NAIVE_BAYES Fold 6 Accuracy: 81.91%

NAIVE_BAYES Fold 6 Confusion Matrix:

	Bad	Good	Very Good
Bad	1477	140	11
Good	181	257	24
Very Good	4	26	14

NAIVE_BAYES Fold 7 Accuracy: 80.74%

NAIVE_BAYES Fold 7 Confusion Matrix:

	Bad	Good	Very Good
Bad	1466	164	10
Good	198	238	20
Very Good	1	18	19

NAIVE_BAYES Fold 8 Accuracy: 81.30%

NAIVE_BAYES Fold 8 Confusion Matrix:

	Bad	Good	Very Good
Bad	1466	128	12
Good	220	248	22
Very Good	1	16	21

NAIVE_BAYES Fold 9 Accuracy: 80.69%

NAIVE_BAYES Fold 9 Confusion Matrix:

	Bad	Good	Very Good
Bad	1454	149	16

Good	194	248	28
Very Good	0	25	20

NAIVE_BAYES Fold 10 Accuracy: 81.77%

NAIVE_BAYES Fold 10 Confusion Matrix:

	Bad	Good	Very Good
Bad	1469	161	5
Good	171	260	30
Very Good	2	20	16

NAIVE_BAYES Confusion Matrix:

	Bad	Good	Very Good
Bad	14746	1399	83
Good	1926	2538	258
Very Good	22	190	183

NAIVE_BAYES Sample Predictions (Title, Actual, Predicted, Correct):

Title: Four Seasons: Fuuka to Nanami, Actual: Bad, Predicted: Bad, Correct: True

Title: Meitantei Conan: Keisatsu Gakkou-hen - Wild Police Story CASE. Furuya Rei, Actual: Bad, Predicted: Bad, Correct: True

Title: Digimon Frontier: Ornismon Fukkatsu!!, Actual: Bad, Predicted: Bad, Correct: True

Title: Pocket Monsters Best Wishes!: Victini to Shiroki Eiyuu Reshiram, Actual: Bad, Predicted: Bad, Correct: True

Title: Sabagebu! OVA SPECIAL MISSION, Actual: Good, Predicted: Good, Correct: True

Title: Chihayafuru 2: Waga Mi Yo ni Furu Nagame Seshi Ma ni, Actual: Good, Predicted: Good, Correct: True

Title: No Guns Life 2, Actual: Good, Predicted: Good, Correct: True

Title: Cuticle Tantei Inaba, Actual: Good, Predicted: Good, Correct: True

Title: Sasaki to Miyano, Actual: Very Good, Predicted: Good, Correct: False

Title: Made in Abyss: Fukaki Tamashii no Reimei, Actual: Very Good, Predicted: Very Good, Correct: True

NAIVE_BAYES Evaluation Metrics (10-Fold CV):

Accuracy: 81.83%
Precision: 0.81
Recall: 0.82
F1-Score: 0.82
Average Similarity: 0.3510
Coverage: 1.0000
Diversity: 4.2804

6. Hasil akurasi KNN dan Naive Bayes jika anime dipisah <70 jelek, 70-80 bagus, >=80 sangat bagus

KNN Recommendations for: shingeki no kyojin

KNN - Mencari anime termirip dengan: Shingeki no Kyojin (kecocokan: 100%)

	Recommended Anime	Similarity	Rating
0	Shingeki no Kyojin Season 2	27.88%	84.0
1	Shingeki no Kyojin: Chronicle	25.24%	76.0
2	Shingeki no Kyojin Season 3 Part 2	24.94%	89.0
3	Shingeki no Kyojin OVA	24.11%	76.0
4	Shingeki no Kyojin Zenpen: Guren no Yumiya	23.63%	75.0

Tags	Top Genres	Top
0 Action, Drama, Fantasy Shapeshifting	Survival, Post-Apocalyptic,	
1 Action, Drama, Fantasy Tragedy	Kaiju, Post-Apocalyptic,	
2 Action, Drama, Fantasy Conspiracy	Dystopian, Military,	
3 Action, Drama, Fantasy Power	Kaiju, Military, Super	
4 Action, Drama, Fantasy Kaiju	Super Power, Medieval,	

Naive Bayes Recommendations for: shingeki no kyojin

Naive Bayes - Mencari anime termirip dengan: Shingeki no Kyojin (kecocokan: 100%)

Recommended Anime Similarity Rating \

0	Shingeki no Kyojin Season 2	100.00%	84.0
1	Shingeki no Kyojin Season 3 Part 2	93.00%	89.0
2	Shingeki no Kyojin: Chronicle	89.40%	76.0
3	Shingeki no Kyojin: The Final Season	89.08%	86.0
4	Shingeki no Kyojin Season 3	88.26%	85.0

Top Genres

Top

Tags

0	Action, Drama, Fantasy	Survival, Post-Apocalyptic, Shapeshifting
1	Action, Drama, Fantasy	Dystopian, Military, Conspiracy
2	Action, Drama, Fantasy	Kaiju, Post-Apocalyptic, Tragedy
3	Action, Drama, Fantasy	War, Tragedy, Military
4	Action, Drama, Fantasy	Conspiracy, Military, Tragedy

Evaluating KNN...

Best k for KNN: 6 (Average Accuracy: 89.15%)

KNN Fold 1 Accuracy: 88.99%

KNN Fold 1 Confusion Matrix:

	Bad	Good	Very Good
Bad	1829	23	4
Good	177	63	3
Very Good	15	13	8

KNN Fold 2 Accuracy: 89.13%

KNN Fold 2 Confusion Matrix:

	Bad	Good	Very Good
Bad	1840	25	1
Good	172	60	4
Very Good	18	12	3

KNN Fold 3 Accuracy: 89.51%

KNN Fold 3 Confusion Matrix:

	Bad	Good	Very Good
Bad	1835	25	0
Good	159	71	3
Very Good	25	12	5

KNN Fold 4 Accuracy: 88.90%

KNN Fold 4 Confusion Matrix:

	Bad	Good	Very Good
Bad	1826	19	1
Good	190	63	2
Very Good	16	9	9

KNN Fold 5 Accuracy: 88.15%

KNN Fold 5 Confusion Matrix:

	Bad	Good	Very Good
Bad	1826	25	0
Good	181	51	5
Very Good	27	15	5

KNN Fold 6 Accuracy: 89.55%

KNN Fold 6 Confusion Matrix:

	Bad	Good	Very Good
Bad	1841	21	3
Good	158	63	4
Very Good	22	15	7

KNN Fold 7 Accuracy: 89.74%

KNN Fold 7 Confusion Matrix:

	Bad	Good	Very Good
Bad	1837	24	1
Good	162	69	3
Very Good	12	17	9

KNN Fold 8 Accuracy: 89.22%

KNN Fold 8 Confusion Matrix:

	Bad	Good	Very Good
Bad	1835	13	0
Good	190	57	1
Very Good	14	12	12

KNN Fold 9 Accuracy: 89.32%

KNN Fold 9 Confusion Matrix:

	Bad	Good	Very Good
Bad	1834	21	0
Good	167	65	2
Very Good	16	22	7

KNN Fold 10 Accuracy: 88.94%

KNN Fold 10 Confusion Matrix:

	Bad	Good	Very Good
Bad	1832	27	0
Good	174	61	2
Very Good	24	9	5

KNN Confusion Matrix:

	Bad	Good	Very Good
Bad	18335	223	10
Good	1730	623	29
Very Good	189	136	70

KNN Sample Predictions (Title, Actual, Predicted, Correct):

Title: Youki ni Utaeba, Actual: Bad, Predicted: Bad, Correct: True

Title: Mao Yu Shu, Actual: Bad, Predicted: Bad, Correct: True

Title: Kanamemo, Actual: Bad, Predicted: Bad, Correct: True

Title: Xuese Cangqiong, Actual: Bad, Predicted: Bad, Correct: True

Title: Maria-sama ga Miteru 3rd Season OVA, Actual: Good, Predicted: Bad, Correct: False

Title: Durarara!!x2 Ten: Onoroke Chakapoko, Actual: Good, Predicted: Bad, Correct: False

Title: 30-sai made Doutei da to Mahou Tsukai ni Nareru Rashii, Actual: Good, Predicted: Bad, Correct: False

Title: Ichigo Mashimaro OVA, Actual: Good, Predicted: Good, Correct: True

Title: Major: World Series, Actual: Very Good, Predicted: Very Good, Correct: True

Title: Shoujo☆Kageki Revue Starlight Movie, Actual: Very Good, Predicted: Good, Correct: False

KNN Evaluation Metrics (10-Fold CV):

Accuracy: 89.14%

Precision: 0.87

Recall: 0.89

F1-Score: 0.87

Average Similarity: 0.5733

Coverage: 0.9411

Diversity: 2.8197

Evaluating Naive Bayes...

NAIVE_BAYES Fold 1 Accuracy: 86.56%

NAIVE_BAYES Fold 1 Confusion Matrix:

	Bad	Good	Very Good
Bad	1723	120	13
Good	112	112	19
Very Good	5	18	13

NAIVE_BAYES Fold 2 Accuracy: 87.03%

NAIVE_BAYES Fold 2 Confusion Matrix:

	Bad	Good	Very Good
Bad	1737	119	10
Good	102	108	26
Very Good	5	15	13

NAIVE_BAYES Fold 3 Accuracy: 86.51%

NAIVE_BAYES Fold 3 Confusion Matrix:

	Bad	Good	Very Good
Bad	1707	135	18
Good	94	119	20
Very Good	5	16	21

NAIVE_BAYES Fold 4 Accuracy: 88.01%

NAIVE_BAYES Fold 4 Confusion Matrix:

	Bad	Good	Very Good
Bad	1738	99	9
Good	110	126	19
Very Good	4	15	15

NAIVE_BAYES Fold 5 Accuracy: 87.82%

NAIVE_BAYES Fold 5 Confusion Matrix:

	Bad	Good	Very Good
Bad	1731	103	17
Good	97	119	21
Very Good	2	20	25

NAIVE_BAYES Fold 6 Accuracy: 85.47%

NAIVE_BAYES Fold 6 Confusion Matrix:

	Bad	Good	Very Good
Bad	1718	127	20
Good	113	94	18
Very Good	5	27	12

NAIVE_BAYES Fold 7 Accuracy: 87.25%

NAIVE_BAYES Fold 7 Confusion Matrix:

	Bad	Good	Very Good
Bad	1726	119	17
Good	104	116	14
Very Good	2	16	20

NAIVE_BAYES Fold 8 Accuracy: 87.39%

NAIVE_BAYES Fold 8 Confusion Matrix:

	Bad	Good	Very Good
Bad	1719	109	20
Good	106	128	14
Very Good	5	15	18

NAIVE_BAYES Fold 9 Accuracy: 86.93%

NAIVE_BAYES Fold 9 Confusion Matrix:

	Bad	Good	Very Good
Bad	1711	120	24
Good	89	126	19
Very Good	2	25	18

NAIVE_BAYES Fold 10 Accuracy: 87.63%

NAIVE_BAYES Fold 10 Confusion Matrix:

	Bad	Good	Very Good
Bad	1734	114	11

Good	95	121	21
Very Good	2	21	15

NAIVE_BAYES Confusion Matrix:

	Bad	Good	Very Good
Bad	17244	1165	159
Good	1022	1169	191
Very Good	37	188	170

NAIVE_BAYES Sample Predictions (Title, Actual, Predicted, Correct):

Title: Youki ni Utaeba, Actual: Bad, Predicted: Bad, Correct: True

Title: Mao Yu Shu, Actual: Bad, Predicted: Bad, Correct: True

Title: Kanamemo, Actual: Bad, Predicted: Bad, Correct: True

Title: Xuese Cangqiong, Actual: Bad, Predicted: Bad, Correct: True

Title: Maria-sama ga Miteru 3rd Season OVA, Actual: Good, Predicted: Bad, Correct: False

Title: Durarara!!x2 Ten: Onoroke Chakapoko, Actual: Good, Predicted: Bad, Correct: False

Title: 30-sai made Doutei da to Mahou Tsukai ni Nareru Rashii, Actual: Good, Predicted: Good, Correct: True

Title: Ichigo Mashimaro OVA, Actual: Good, Predicted: Bad, Correct: False

Title: Major: World Series, Actual: Very Good, Predicted: Bad, Correct: False

Title: Shoujo★Kageki Revue Starlight Movie, Actual: Very Good, Predicted: Good, Correct: False

NAIVE_BAYES Evaluation Metrics (10-Fold CV):

Accuracy: 87.06%

Precision: 0.88

Recall: 0.87

F1-Score: 0.87

Average Similarity: 0.3510

Coverage: 1.0000

Diversity: 4.5063

7. Koding Sentimen Analisis (Google Colab)

https://colab.research.google.com/drive/1ul-iB5HpBq-mwAGjIC_UyR9ebmZOhddy#scrollTo=MkUjnZRjSv2g

8. Koding API Anilist, KNN, dan Pembuatan Website (Google Colab)

https://colab.research.google.com/drive/13q0B07S97_6_3NX5SQ7EZDlQuDTTzqHY#scrollTo=Q9U6dpceHn4F



UNIVERSITAS
MA CHUNG