PERBANDINGAN METODE *NAÏVE BAYES CLASSIFIER* DAN *RANDOM FOREST* UNTUK KLASIFIKASI FILM BERDASARKAN USIA PENONTON

Diajukan Sebagai Syarat Untuk Menyelesaikan
Pendidikan Program Strata-1 Pada
Jurusan Teknik Informatika Fakultas Ilmu Komputer UNSRI

Oleh :

MSY RIZKIA YUNIANDARI
NIM : 09021282126096

Jurusan Teknik Informatika
FAKULTAS ILMU KOMPUTER UNIVERSITAS SRIWIJAYA
2025

HALAMAN PENGESAHAN

SKRIPSI

Perbandingan Metode Naive Bayes Classifier dan Random Forest untuk Klasifikasi Film Berdasarkan Usia Penonton

Sebagai salah satu syarat untuk penyelesaian studi di

Program Studi S1 Teknik Informatika

Oleh:

MSY RIZKIA YUNIANDARI

09021282126096

Pembimbing 1 : **Alvi Syahrini Utami, M.Kom.**

NIP. 197812222006042003

Mengetahui

Ketua Jurusan Teknik Informatika

Hadipurnawan Satria, Ph.D

198004182020121001

TANDA LULUS UJIAN KOMPREHENSIF

Pada hari Jumat tanggal 13 Juni 2025 telah dilaksanakan ujian komprehensif skripsi oleh Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya

Nama : Msy Rizkia Yuniandari
Nim : 09021282126096
Judul : Perbandingan Metode *Naive Bayes Classifier* dan *Random Forest* untuk Klasifikasi Film Berdasarkan Usia Penonton.

dan dinyatakan **LULUS.**

1. Ketua Penguji

   Endang Lestari, S.Kom , M.T
   NIP. 197811172006042001 .....

2. Penguji I

   M. Qurhanul Rizqie, M.T., Ph.D
   NIP. 198712032022031006 .....

3. Pembimbing I

   Alvi Syahrini Utami, M.Kom
   NIP. 197812222006042003 .....

Mengetahui
Ketua Jurusan Teknik Informatika

Hadipurnawan Satria, Ph.D.
NIP. 198004182020121001

HALAMAN PERNYATAAN

Yang bertanda tangan di bawah ini :

Nama : Msy Rizkia Yuniandari

NIM : 09021282126096

Program Studi : Teknik Informatika

Judul Skripsi : Perbandingan Metode *Naïve Bayes Classifier* dan *Random Forest* untuk Klasifikasi Film Berdasarkan Usia Penonton

Hasil pengecekan *Software* Turnitin : 5%

Menyatakan bahwa laporan tugas akhir saya merupakan hasil karya saya sendiri dan bukan hasil penjiplakan/plagiat. Apabila ditemukan unsur penjiplakan/plagiat dalam laporan proyek ini, maka saya bersedia menerima sanksi akademik dari Universitas Sriwijaya sesuai dengan ketentuan yang berlaku.

Demikian pernyataan ini saya buat dengan sebenarnya dan tidak ada paksaan dari pihak mana pun.

Inderalaya, 16 Juni 2025

Penulis,

Msy Rizkia Yuniandari

NIM. 09021282126096

MOTTO DAN PERSEMBAHAN

F-E-A-R has two meanings 'Forget Everything and Run' of 'Face Everything and Rise.' The choice is yours.

- Zid Ziglar

Kupersembahkan karya tulis ini kepada:

- ❖ Orang Tua dan Saudaraku
- ❖ Teman-teman Seperjuangan
- ❖ Fakultas Ilmu Komputer
- ❖ Universitas Sriwijaya

ABSTRACT

The rapid growth of the film industry in recent years has created a need for an accurate age classification system to ensure that content is appropriate to its target audience. One approach that can be used to identify the age classification of a film is through analysis of the subtitle text, without relying on visual or audio elements. This study aims to compare the performance of two classification algorithms, namely Naïve Bayes Classifier and Random Forest, in categorizing films based on audience age groups. The classification process involves a series of text preprocessing steps and feature extraction using TF-IDF method. The evaluation results show that Random Forest achieved an average accuracy of 70,83% while Naïve Bayes obtained an average accuracy of 68,75%. Although the difference in accuracy is not substantial, it can be concluded that Random Forest perfoms better and is more suitable for film classification based on age categories.

Keywords: *movie classification, Subtitle, TF-IDF, Naïve Bayes Classifier, Random Forest*

ABSTRAK

Perkembangan industri perfilman yang pesat dalam beberapa tahun terakhir menimbulkan kebutuhan akan sistem klasifikasi usia yang akurat untuk menjaga kesesuaian konten dengan kelompok penontonnya. Salah satu pendekatan yang dapat digunakan untuk mengidentifikasi klasifikasi usia film adalah melalui analisis teks *subtitle*, tanpa bergantung pada elemen visual atau audio. Penelitian ini bertujuan untuk membandingkan kinerja dua algoritma klasifikasi, yaitu *Naïve Bayes Classifier* dan *Random Forest*, dalam mengelompokkan film berdasarkan kategori usia penonton. Proses klasifikasi dilakukan dengan menerapkan tahapan pra-pemrosesan teks dan ekstraksi fitur menggunakan TF-IDF. Hasil evaluasi menunjukkan bahwa *Random Forest* mampu mencapai akurasi sebesar 70,83%, sementara *Naïve Bayes Classifier* memperoleh akurasi sebesar 68,75%. Meskipun selisih akurasinya tidak terlalu signifikan, dapat disimpulkan bahwa *Random Forest* memiliki performa yang lebih unggul dan lebih sesuai digunakan dalam klasifikasi film berdasarkan kategori usia penonton.

Kata Kunci: *klasifikasi film, Subtitle, TF-IDF, Naïve Bayes Classifier, Random Forest*

KATA PENGANTAR

Segala puji dan syukur penulis panjatkan kehadirat Allah SWT atas berkat dan rahmat-Nya yang telah diberikan kepada Penulis sehingga dapat menyelesaikan Tugas Akhir ini yang berjudul "Perbandingan Metode *Naïve Bayes Classifier* dan *Random Forest* untuk Klasifikasi Film Berdasarkan Usia Penonton". Tugas akhir ini disusun untuk memenuhi salah satu syarat untuk menyelesaikan pendidikan program Strata-1 pada Fakultas Ilmu Komputer Program Studi Teknik Informatika di Universitas Sriwijaya.

Dalam menyelesaikan tugas akhir ini, banyak pihak yang telah memberikan bantuan dan dukungan baik secara langsung maupun tidak secara langsung. Untuk itu, Penulis ingin menyampaikan rasa terima kasih kepada:

1. Bapak Prof. DR. Erwin, S.Si., M.Si. , selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya.
2. Bapak Hadipurnawan Satria, M.Sc.,Ph.D. selaku Ketua Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.
3. Ibu Alvi Syahrini Utami, M.Kom. selaku dosen pembimbing yang telah memberikan bimbingan, masukan, serta motivasi dalam menyelesaikan skripsi ini.
4. Seluruh dosen di Program Studi Teknik Informatika yang telah memberikan ilmu selama masa perkuliahan.
5. Kepada kedua orang tuaku tersayang, papa dan mama, panutan dalam hidupku. Terima kasih telah menjadi sumber kekuatan dan support system dalam setiap

langkahku. Terima kasih atas setiap pengorbanan dan kerja keras untuk memberikan yang terbaik kepada penulis, atas doa-doa yang tak pernah putus, dan kasih sayang yang tak pernah berkurang hingga akhirnya penulis dapat menyelesaikan pendidikan dan meraih gelar sarjana.

6. Keempat saudara penulis, mangepi, cek eli, kakcik dan macik yang selalu memberikan dukungan, memotivasi dan mendoakan penulis.
7. Eka selaku teman sekaligus sahabat penulis yang senantiasa memberi semangat dan selalu hadir dalam setiap proses selama perkuliahan hingga penyusunan skripsi ini. Terima kasih atas setiap waktu yang telah kamu luangkan untuk mendengar keluh kesah penulis hingga penyusunan skripsi ini selesai.
8. Kepada Cici dan Naufal terima kasih telah mendengarkan keluh kesah penulis, berkontribusi dalam penulisan skripsi ini, dan selalu memberikan semangat.
9. Fourth Nattawat yang turut memberikan andil secara tidak langsung lewat energi positifnya, sehingga mampu membuat hari-hari selama proses pengerjaan tugas akhir ini menjadi lebih berwarna, penuh semangat dan motivasi.
10. Teman-teman TIREG L3 2021 dan seluruh teman-teman Teknik Informatika Universitas Sriwijaya.
11. Semua pihak yang telah membantu dalam penyusunan tugas akhir ini yang tidak dapat disebutkan satu persatu.
12. Terakhir terima kasih untuk diriku sendiri. Semoga ini jadi pengingat bahwa kamu mampu, dan layak untuk bangga pada dirimu sendiri. Berbahagialah

selalu dimanapun berada. Adapun lebih dan kurangnya mari merayakan diri sendiri.

Penulis menyadari dalam penyusunan Tugas Akhir ini masih terdapat banyak kekurangan yang disebabkan keterbatasan pengetahuan dan pengalaman. Oleh karena itu, kritik dan saran yang membangun sangat diharapkan. Akhir kata, dengan segala kerendahan hati, semoga Tugas Akhir ini dapat berguna dan bermanfaat bagi kita semua.

Inderalaya, 13 Juni 2025

Msy Rizkia Yuniandari

DAFTAR ISI

DAFTAR GAMBAR

DAFTAR TABEL

BAB I

PENDAHULUAN

1.1 Pendahuluan

Pada bab pendahuluan ini akan membahas tentang latar belakang diambilnya topik "Perbandingan Metode *Naïve Bayes Classifier* dan *Random Forest* untuk Klasifikasi Film Berdasarkan Usia Penonton". Pada bab ini juga berisi penjelasan dan alasan mengenai penelitian yang dimulai dari latar belakang masalah, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah, dan sistematika penulisan laporan tugas akhir.

1.2 Latar Belakang

Dalam era digital saat ini, akses terhadap film semakin mudah. Namun, dengan berbagai jenis film yang tersedia, penting untuk memahami dan mengklasifikasikan film berdasarkan usia penonton. Klasifikasi ini tidak hanya membantu penonton dalam memilih konten yang sesuai, tetapi juga penting bagi orang tua dalam mengawasi apa yang ditonton oleh anak-anak mereka dan juga panduan untuk memilih film yang akan ditonton oleh anaknya (Agung et al., 2020). Klasifikasi film ini tidak menunjukkan baik atau buruknya suatu film dari segi kualitas, melainkan hanya memberikan panduan tentang konten-konten tidak pantas untuk umur tertentu.

Menurut *Communication Research Reports*, menemukan bahwa 68% film PG-13 mengandung kekerasan seksual yang tidak tercermin dalam *rating* nya (Leone & Houle, 2019). Sehubungan dengan hal tersebut, ketidaksesuaian antara

konten film dan *rating* usia yang diberikan menimbulkan kekhawatiran serius tentang efektivitas sistem *rating* dalam melindungi penonton muda.

Sistem sertifikat *rating* berdasarkan usia penonton akan memberikan label terhadap sebuah film yang menggambarkan elemen-elemen kontennya dan dinilai berdasarkan suatu pedoman tertentu oleh badan yang memiliki kewenangan dalam penggelompokkan batas usia dalam film. Dalam hal ini, sertifikasi rating sistem *Motion Picture Assoiation of America* (MPAA) telah digunakan sebagai standar klasifikasi film. Seiring dengan meningkatnya produksi film, dibutuhkan metode otomatis untuk mengklasifikasikan film berdasarkan konten dan tema. Pendekatan yang dapat digunakan adalah pemrosesan bahasa alami (*Natural Language Processing/NLP*) dan *machine learning.*

Penelitian ini menggunakan metode *Naïve Bayes Classifier* dan *Random Forest* dalam klasifikasi film berdasarkan usia penonton. Menurut Muhammad et al. (2017), *Naïve Bayes Classifier* adalah metode klasifikasi statistik berdasarkan *teorema bayes* yang dapat digunakan untuk memprediksi probabilitas kelas dalam data. Klasifikasi menggunakan *Naive Bayes* telah terbukti sangat akurat dan cepat ketika diterapkan pada dataset besar. Keuntungan klasifikasi adalah membutuhkan sangat sedikit data pelatihan untuk memperkirakan parameter yang diperlukan untuk klasifikasi (rata-rata dan varians variabel). Sementara *Random Forest* merupakan metode *ensemble learning* yang menggabungkan *multiple decision trees* untuk menghasilkan prediksi yang lebih stabil dan akurat. Keuntungan dari penggunaan kedua metode ini adalah kemampuannya dalam menangani dataset

besar dengan komputasi yang efisien, serta kemampuan untuk mengurangi *overfitting* melalui pendekatan yang berbeda.

Pada penelitian sebelumnya, yang dilakukan oleh Kirsehir Ahi Evran dengan judul "*A Smart Movie Suitability Rating System Based on Subtitle*" (2022). Diketahui peneliti menggunakan berbagai metode machine learning untuk klasifikasi film berdasarkan usia penonton. Peneliti mengevaluasi performa dengan membandingkan beberapa metode *machine learning* yaitu *Random Forest, Support Vector Machine,* dan *K-Nearest Neighbour.* Hasil evaluasi menunjukkan bahwa metode *Random Forest* memberikan kinerja terbaik dengan tingkat akurasi mencapai 84%. Metode lain seperti, *Support Vector Machine* menunjukkan performa yang cukup baik dengan akurasi 76%. Sementara itu, pada *K-Nearest Neighbour* menunjukkan kinerja yang relatif lebih rendah dengan akurasi 45% . Perbedaan signifikan dalam tingkat akurasi ini menunjukkan bahwa Random Forest memiliki kemampuan yang lebih baik dalam menangani kompleksitas klasifikasi film berdasarkan usia penonton.

Selain itu juga, dari penelitian Rifqy Rosyidah Ilmi dkk (2023) melakukan penelitian nilai *rating* film IMDb dengan menggunakan *Decision Tree*. Dataset yang digunakan terdapat 28 atribut diantaranya, judul film, durasi, *content rating* (G, PG, PG-13 dan R), sutradara, aktor dan artis dan banyak lagi. Hasil penelitian menunjukkan *decision tree* mampu melakukan prediksi nilai rating film IMDb dan menunjukkan faktor yang mempengaruhi tinggi atau rendah nilai *rating* film. Nilai akurasi yang didapat pada data *training*, validasi dan *testing* berturut – turut adalah 75%, 72% dan 70%.

Berdasarkan uraian diatas, penelitian ini akan membandingkan metode *Naïve Bayes Classifier* dengan metode *Random Forest* untuk klasifikasi film berdasarkan usia penonton, hasil yang diharapkan pada penelitian ini dapat mengetahui metode yang paling sesuai dalam mengklasifikasi.

1.3 Rumusan Masalah

Berdasarkan latar belakang diatas, maka penulis mendapatkan suatu rumusan masalah yang menjadi dasar penelitian tugas akhir ini yaitu, sebagai berikut:

1. Bagaimana cara mengklasifikasikan film berdasarkan usia penonton menggunakan metode *Naïve Bayes Classifier* dan *Random Forest*?
2. Berapa besar tingkat akurasi yang dihasilkan dari kedua metode tersebut dalam mengklasifikasikan film berdasarkan usia penonton?

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah diatas, tujuan dilakukannya penelitian ini yaitu sebagai berikut:

1. Mengembangkan perangkat lunak menggunakan metode *Naïve Bayes Classifier* dan *Random Forest* untuk mengklasifikasikan film berdasarkan usia penonton.
2. Mengukur seberapa besar tingkat akurasi yang dihasilkan untuk mengetahui metode manakah yang paling sesuai dalam mengklasifikasikan film berdasarkan usia penonton.

1.5 Manfaat Penelitian

Manfaat yang dapat diperoleh dari penelitian ini adalah sebagai berikut:

1. Dapat mengetahui batasan umur di setiap film berdasarkan *input file subtitle* dari film yang diuji.
2. Sebagai sarana untuk panduan orang tua dalam pemilihan konten film yang sesuai sehingga bisa dijadikan arahan untuk memberikan tontonan kepada anak-anak.
3. Dapat mengetahui metode yang paling sesuai untuk klasifikasi film berdasarkan usia penonton dengan menghitung dari nilai akurasi keduanya.
4. Hasil penelitian dapat dijadikan sebagai sumber atau referensi bagi peneliti yang ingin membahas tentang klasifikasi film berdasarkan usia penonton menggunakan metode *Naïve Bayes Classifier* dan *Random Forest.*

1.6 Batasan Masalah

Batasan masalah pada penelitian ini adalah sebagai berikut:

1. *File subtitle* menggunakan bahasa inggris.
2. *File subtitle* diambil dari website penyedia subtitle dengan jenis file .srt atau disebut dengan softsub.
3. *Output* rating hanya sebatas G, PG, PG-13, dan R yang telah ditetapkan IMDb[1].
4. Terdapat 240 buah dataset subtitle yang terdiri dari 4 macam *rating* yaitu G, PG, PG-13, dan R yang masing-masing berjumlah 60 buah.

[1] http://imdb.com/?ref_=nv_home

1.7 Sistematika Penulisan

Sistematika penulisan tugas akhir ini mengikuti standar penulisan tugas akhir Fakultas Ilmu Komputer Universitas Sriwijaya yaitu sebagai berikut:

BAB I. PENDAHULUAN

Pada bab ini menjelaskan pendahuluan mengenai latar belakang, rumusan masalah, tujuan, dan manfaat penelitian, batasan masalah dan sistematika penulisan dalam penelitian.

BAB II. TINJAUAN PUSTAKA

Pada bab ini, akan dijelaskan mengenai dasar-dasar teori yang berhubungan dengan penelitian: seperti, metode *Naïve Bayes Classifier*, *Random Forest* dan kajian literatur yang relevan dengan penelitian ini.

BAB III. METODE PENELITIAN

Pada bab ini akan dibahas mengenai tahapan yang akan dilakukan pada penelitian ini, seperti, metode pengumpulan data dan kerangka kerja penelitian, masing masing kriteria penelitian yang digunakan dijelaskan dengan rinci. Di akhir bab ini berisi perancangan manajemen proyek pada pelaksanaan penelitian.

BAB IV. PENGEMBANGAN PERANGKAT LUNAK

Pada bab ini akan membahas tentang perancangan dan implementasi perangkat lunak klasifikasi film berdasarkan usia penonton yang telah dikembangkan.

BAB V. HASIL DAN ANALISIS PENELITIAN

Pada bab ini akan menguraikan hasil dan analisa hasil dari pengujian perangkat lunak secara keseluruhan.

BAB VI. KESIMPULAN DAN SARAN

Pada bab ini berisi kesimpulan dari semua uraian-uraian pada bab-bab sebelumnya dan juga berisi saran-saran yang diharapkan berguna dalam penelitian selanjutnya.

1.8 Kesimpulan

Pada bab I ini telah dibahas mengenai penelitian yang akan dilaksanakan yaitu perbandingan *Naïve Bayes Classifier* dan *Random Forest* dalam mengklasifikasikan film berdasarkan usia penonton.

DAFTAR PUSTAKA

Razaq, M.T, N. Dede, dan N. Hani. 2023. Analisis Sentimen Review Film Menggunakan Naïve Bayes Classifier dengan Fitur TF-IDF. (PP 1698-1710).

Yang, L. and Park, J. 2020. "Subtitle Analysis in Multimedia Translation". International Journal of Media Studies, 18(2), 45-62.

Chen, W., et al. (2020). "SubRip Subtitle Format: Technical Specifications". International Journal of Media Conversion, 15(2), 112-128.

Thompson, R. (2020). "Global Film Rating Systems". Media Regulation Journal, 15(2), 78-95.

MPAA. 2024. Ratings Guide Mean. (https://www.showbizjunkies.com/mpaa-ratings/, diakses 2 Februari 2025)

LSF. Lembaga Sensor Film. (https://www.lsf.go.id/wp-content/uploads/2021/03/PP-LSF.pdf, diakses 4 Februari 2025)

Jurafsky, D and Martin, J. H. 2020. Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition. Upper Saddle River: Prentice Hall.

Shafirra, N. A., & Irhamah, I. (2020). Klasifikasi Sentimen Ulasan Film Indonesia dengan Konversi Speech-to-Text (STT) Menggunakan Metode Convolutional Neural Network (CNN). Jurnal Sains Dan Seni ITS, 9(1). https://doi.org/10.12962/j23373520.v9i1.51825.

Young, T., Hazarika, D., Poria, S., & Cambria, E. (2018). Recent trends in deep learning based natural language processing. IEEE Computational intelligence magazine, 13(3), 55-75.

McCallum, A., & Nigam, K. (1998). "A comparison of event models for naive bayes text classification." AAAI-98 workshop on learning for text categorization, 752(1), 41-48.

Read, J., Pfahringer, B., Holmes, G., & Frank, E. (2011). Classifier Chains For Multi-Label Classification. Machine Learning, 85(3), 333–359. Https://Doi.Org/10.1007/S10994-011-5256-5.

Aurora, Rahul. (2023). Advantages and Disadvantages of Naïve Bayes Classifier. Retrieved from https://iq.opengenus.org/advantages-and-disadvantages-of-naive-bayes-algorithm/

Chen, X., Wang, Y., & Liu, Z. (2022). Automatic film rating prediction using subtitle analysis and ensemble methods. Journal of Machine Learning Research, 23(1), 1-34.

Zhang, Y., & Li, H. (2023). "Advanced Text Classification Techniques Using Enhanced TF-IDF and Machine Learning Algorithms." *Journal of Artificial Intelligence Research*, 45(3), 215-237.

Tsoumakas, G., & Katakis, I. (2007). Multi-Label Classification: An Overview. International Journal Of Data Warehousing And Mining, 3(3), 1–13. Https://Doi.Org/10.4018/Jdwm.2007070101

Young, T., Hazarika, D., Poria, S., & Cambria, E. (2018). Recent Trends In Deep Learning Based Natural Language Processing [Review Article]. In IEEE Computational Intelligence Magazine (Vol. 13, Issue 3, Pp. 55–75). Institute Of Electrical And Electronics Engineers Inc. Https://Doi.Org/10.1109/MCI.2018.2840738

Hanif, Jumly Ahmad., Farid, Mifta & Hasanah, Barokatun. (2023). Penerapan Natural language Processing untuk Klasifikasi Bidang Minat, 41-49.

Naeem, Muhammad Zaid., Rustam, Furqan., Mehmood, Arif., Ashraf, Imran., Muiizuddin., Gyu, Sang Choi. (2022). *Classification of movie reviews using term frequency-inverse document frequency and optimized machine learning algorithms*.

Tripathi S, Mehrotra R, Bansal V, Upadhyay S. (2020). Analyzing Sentiment using IMDb dataset.

Nafis NSM, Awang S. 2021. An enchanced hybrid feature selection technique using term frequency-inverse document frequency and support vector machine-recursive feature elimination for sentiment classification. *IEEE Access* **9**:52177-52192

Mitta, Vanishka., Guru, P. M. S., Vishwakarma, Harsh Kumar., Ganesh D. R., Chandrappa S. (2022). Sentimental Analysis of Movie Reviw Based on Naïve Bayes and Random Forest Technique.

M. Hidayat, R. Hidayat and D. O. Kurniawati, "Comparison of the use of bigrams and stopword removal for classification using naive bayes (case study on sentiment analysis of by. u internet users)", *2021 International Conference on Software Engineering & Computer Systems and 4th International Conference on Computational Science and Information Management (ICSECS-ICOCSIM)*, pp. 447-452, 2021.

Sandag, G. A. Prediksi Rating Aplikasi App Store Menggunakan Algoritma Random Forest. *Cogito Smart Journal,* 6(2), 2020.

Novenrodumetasaa, Nathania., Raharja, Made Sunia., Suarjaya, Made Agus Dwi. (2023). Analisis Genre Film Berdasarkan Data Subtitle. Open Journal System.

N. K. Rajput and B. A. Grover, “A Multi-label Movie Genre Classification Scheme Based on the Movie’s Subtitle,” Multimed. Tools Appl., vol. 81, no. 22, pp. 32469–32490, 2022, doi: https://doi.org/10.1007/s11042-022-12961-6.

Guia, M., Silva, R. R., & Bernardino, J. (2019). Comparison of Naive Bayes, support vector machine, decision trees and random forest on sentiment analysis. IC3K 2019 - Proceedings of the 11th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management, 525-531. https://doi.org/10.5220/0008364105250531.

Atimi, R. L., & Enda Esyudha Pratama. (2022). Implementasi Model Klasifikasi Sentimen Pada Review Produk Lazada Indonesia. Jurnal Sains Dan Informatika, 8(1), 88–96. https://doi.org/10.34128/jsi.v8i1.419.

Syafarina, G. A., & Zaenuddin. (2023). Implementasi Framework Streamlit Sebagai Prediksi Harga Jual Rumah Dengan Linear Regresi. 7, 2023. Https://Doi.Org/10.47002/Metik.V7i2.680.

LAMPIRAN

Link Github : https://github.com/kiaasthetics/Movie-Classification-Based-on-Subtitle.git