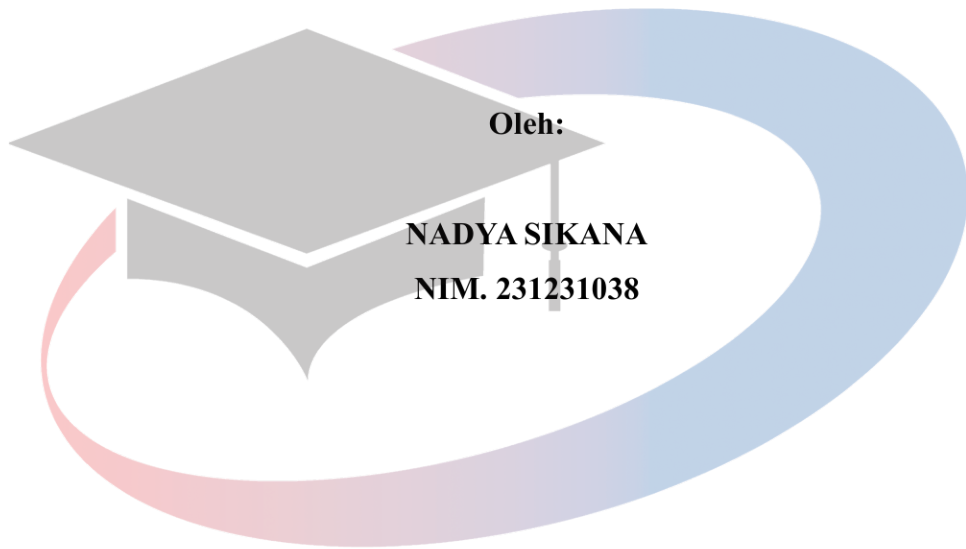


**PREDIKSI KESUKSESAN FILM BERDASARKAN FITUR DAN
KOMENTAR TRAILER MENGGUNAKAN MODEL
ENSEMBLE+LSTM**

TESIS



Oleh:

NADYA SIKANA

NIM. 231231038



PROGRAM STUDI S-2 TEKNOLOGI INFORMASI

FAKULTAS INFORMATIKA

UNIVERSITAS MIKROSKIL

MEDAN

2024

**MOVIE SUCCESS PREDICTION BASED ON FEATURE AND TRAILER
COMMENTS USING ENSEMBLE+LSTM MODEL**

THESIS

By:

NADYA SIKANA

ID NUMBER. 231231038



MAJOR OF S-2 INFORMATION TECHNOLOGY

FACULTY OF INFORMATICS

UNIVERSITAS MIKROSKIL

MEDAN

2024

ABSTRAK

PREDIKSI KESUKSESAN FILM BERDASARKAN FITUR DAN KOMENTAR TRAILER MENGGUNAKAN MODEL ENSEMBLE+LSTM

Abstrak

Memprediksi kesuksesan film menjadi satu aspek yang sangat penting karena terdapat risiko yang tinggi di dalam produksi sebuah film. Tantangannya adalah ketidakpastian di dalam industri film dan memilih model *machine learning* yang tepat. Kita dapat menggabungkan fitur film dan analisis sentimen dari media sosial menggunakan teknik *machine learning* untuk mencapai prediksi kesuksesan film. Metode yang digunakan untuk prediksi berdasarkan fitur film adalah model *Ensemble (Random Forest + Gradient Boosting)*. Sedangkan metode yang digunakan untuk analisis sentimen komentar *trailer* adalah LSTM. Evaluasi model yang digunakan adalah perhitungan RMSE dan akurasi. Hasil akhir dari prediksi kesuksesan mendapatkan RMSE sebesar 0,8807 dan akurasi 91,19%. Ini merupakan peningkatan dari penelitian sebelumnya. Penelitian selanjutnya disarankan untuk menerapkan model pada industri film.

Kata kunci: *analisis sentimen, metode ensemble, prediksi kesuksesan film*

Abstract

Predicting the success of a movie is a very important aspect due to the high risks involved in movie production. The challenge lies in the uncertainty within the movie industry and selecting the appropriate machine learning model. We can combine movie features and sentiment analysis from social media using machine learning techniques to achieve movie success prediction. The methods used for predicting based on movie features are Ensemble models (Random Forest + Gradient Boosting). Meanwhile, the methods used for sentiment analysis of trailer comments is LSTM. The evaluation of the models used is based on RMSE and accuracy calculation. The final prediction of success obtains an RMSE of 0,8807 and an accuracy of 91,19%. This represents an improvement from previous research. Further research is recommended to implement the model in the movie industry.

Keywords: *sentiment analysis, ensemble method, movie success prediction*

KATA PENGANTAR

Ucapan syukur penulis panjatkan kepada Tuhan Yang Maha Esa, atas berkat dan rahmat-Nya penulis dapat menyelesaikan Tesis yang berjudul “PREDIKSI KESUKSESAN FILM BERDASARKAN FITUR DAN KOMENTAR TRAILER MENGGUNAKAN MODEL ENSEMBLE+LSTM”.

Penulis menyampaikan banyak terima kasih kepada:

1. Bapak Dr. Ronsen Purba, M.Sc., selaku Dosen Pembimbing yang telah membimbing dan memberikan saran serta motivasi dalam penyusunan Tesis ini.
2. Bapak Hardy, S.Kom., M.Sc., Ph.D., selaku Rektor Universitas Mikroskil.
3. Bapak Ng Poi Wong, S.Kom., M.T.I., selaku Dekan Fakultas Informatika Universitas Mikroskil.
4. Bapak Ir. Erwin Setiawan Panjaitan, M.M.S.I., Ph.D., selaku Ketua Program Studi S-2 Teknologi Informasi Fakultas Informatika Universitas Mikroskil.
5. Bapak dan Ibu Dosen Program Studi S-2 Teknologi Informasi Universitas Mikroskil yang telah membimbing dan memberikan pengetahuan yang sangat berarti kepada penulis.
6. Orang tua dan keluarga yang senantiasa memberikan dukungan, nasihat dan doa yang selalu dipanjatkan agar ilmu yang penulis dapatkan mampu memberikan manfaat bagi penulis dan orang banyak.
7. Teman-teman dan rekan kerja penulis yang tiada henti dalam memberikan dorongan dan semangat kepada penulis untuk menyelesaikan penulisan tesis ini.

Penulis menyadari bahwa Tesis ini masih terdapat banyak kekurangan karena keterbatasan kemampuan dari penulis. Oleh karena itu, kritik dan saran dari pembaca sangat diharapkan untuk penulisan yang lebih baik di masa yang akan datang dan semoga Tesis ini dapat memberikan manfaat bagi pembaca.

Medan, 09 Juli 2024

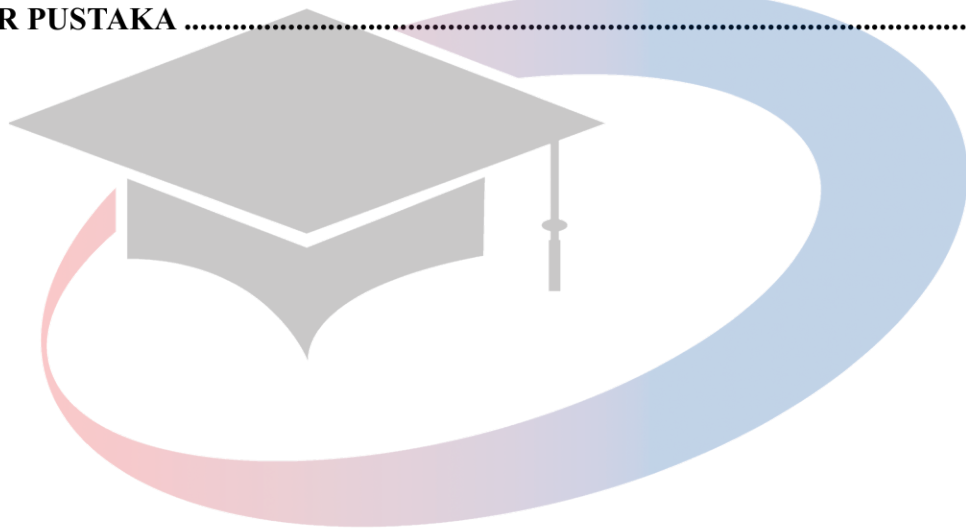
Penulis,

Nadya Sikana

DAFTAR ISI

ABSTRAK	i
KATA PENGANTAR	ii
DAFTAR ISI	iii
DAFTAR GAMBAR	v
DAFTAR TABEL	vi
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Tujuan	2
1.4 Manfaat Penelitian	2
1.5 Ruang Lingkup.....	3
BAB II KAJIAN LITERATUR	4
2.1 Tinjauan Pustaka	4
2.1.1 Prediksi Kesuksesan Film	4
2.1.2 Random Forest	4
2.1.3 Gradient Boosting	5
2.1.4 LSTM (Long-Short Term Memory).....	6
2.1.5 Pengujian Tingkat Keakuratan Hasil Prediksi	7
2.2 Penelitian Terdahulu	8
BAB III METODOLOGI PENELITIAN	10
3.1 Analisis Masalah.....	10
3.2 Data yang Digunakan.....	10
3.3 Metode Penelitian	11
3.4 Teknik Analisis Data	12
BAB IV HASIL DAN PEMBAHASAN	14
4.1 Hasil	14
4.1.1 Membaca Dataset	14
4.1.2 Pembersihan Dataset	15
4.1.3 Pembuatan Model Prediksi Fitur	17
4.1.4 Pembuatan Model Analisis Sentimen Komentar <i>Trailer</i>	17
4.1.5 Prediksi Kesuksesan Film Berdasarkan Fitur dan Komentar <i>Trailer</i>	17
4.1.6 Pemberian Data Baru	18

4.2 Pembahasan.....	20
4.2.1 Perbandingan Prediksi Menggunakan Data Fitur Film-Komentar <i>Trailer</i> dan Data Fitur Film.....	20
4.2.2 Perbandingan Prediksi Menggunakan Data Fitur Film-Komentar <i>Trailer</i> dan Data Komentar <i>Trailer</i>	21
4.2.3 Perbandingan Prediksi Menggunakan Model Ensemble+LSTM dan Model Lainnya	21
BAB V PENUTUP.....	22
5.1 Kesimpulan	22
5.2 Saran	22
DAFTAR PUSTAKA	23



UNIVERSITAS MIKROSKIL

DAFTAR GAMBAR

Gambar 2.1 Alur Kerja <i>Random Forest</i> [14]	5
Gambar 2.2 <i>Pseudocode Gradient Boosting</i> [17]	5
Gambar 2.3 Arsitektur LSTM [20]	7
Gambar 3.1. <i>Dataset Film dan Pemeran</i>	10
Gambar 3.2. <i>Dataset Review Film</i>	11
Gambar 3.3 <i>Flowchart Metode Penelitian</i>	11
Gambar 4.1 <i>Dataset Mentah Film dan Pemeran</i>	14
Gambar 4.2 <i>Dataset Mentah Review Film</i>	15
Gambar 4.3 Tampilan <i>Dataset Film dan Pemeran</i> Setelah Dibersihkan	15
Gambar 4.4 Tampilan <i>Dataset Film dan Pemeran</i> Setelah Dilakukan <i>Encoding</i>	16
Gambar 4.5 Tampilan <i>Dataset Review Film</i> Setelah Dibersihkan	16
Gambar 4.6 Data Gabungan Fitur dan Analisis Sentimen <i>Review Film</i>	18
Gambar 4.7 Data Fitur Film Baru	18
Gambar 4.8 Data Komentar <i>Trailer Film Baru</i>	19
Gambar 4.9 Data Gabungan Film Baru	19
Gambar 4.10 Perbandingan <i>Rating</i> yang Diprediksi dan <i>Rating Asli</i>	20



UNIVERSITAS
MIKROSKIL

DAFTAR TABEL

Tabel 4.1 Perbandingan RMSE dan Akurasi Tiga Model untuk Prediksi Berdasarkan Fitur Film.....	17
Tabel 4.2 Perbandingan Akurasi Model LSTM untuk Analisis Sentimen <i>Review</i> Film.....	17

