

Manual Book

Sistem Klasifikasi Sentimen Ulasan Google Play Store

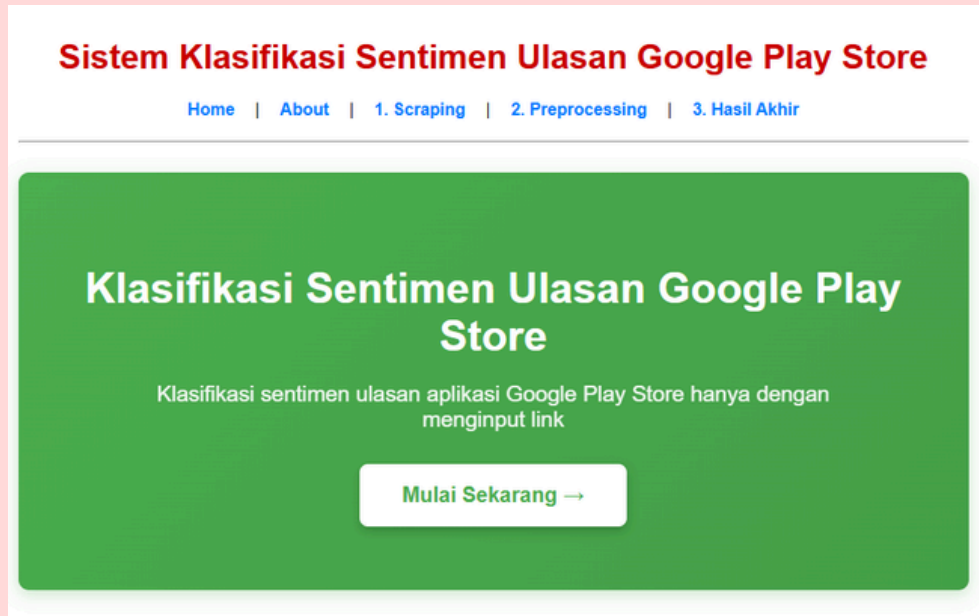


Christ Thaddeus - 535220196

Tentang sistem

Sistem ini adalah aplikasi web untuk klasifikasi sentimen ulasan aplikasi dari Google Play Store. Sistem menggunakan dua model machine learning yang berbeda (XGBoost dan AdaBoost) untuk melakukan prediksi dan membandingkan hasilnya secara bersamaan. Masing-masing model dilatih menggunakan data ulasan aplikasi Gojek di Google Play Store yang telah melalui proses pre-processing.

Halaman Home



Saat mengakses website, halaman yang pertama akan muncul adalah halaman Home untuk mengundang user untuk menggunakan sistem klasifikasi

Halaman About

Tentang Sistem Klasifikasi Sentimen

Sistem ini merupakan aplikasi web untuk klasifikasi sentimen ulasan aplikasi dari Google Play Store. Sistem menggunakan dua model machine learning yang berbeda (XGBoost dan AdaBoost) untuk melakukan prediksi. Masing-masing model dilatih menggunakan data ulasan aplikasi Gojek di Google Play Store dengan hasil akurasi dan F-1 Score sekitar 85% untuk AdaBoost dan 88% untuk XGBoost.

Alur Proses Klasifikasi Sentimen

1 Scraping Data dari Google Play Store

Di tahap awal ini, sistem akan mengambil data ulasan aplikasi langsung dari Google Play Store. Anda hanya perlu memasukkan URL Google Play Store dari aplikasi yang ingin anda lakukan proses klasifikasi sentimen dan menentukan jumlah ulasan yang ingin diambil, maksimal 20,000 ulasan.

2 Preprocessing Data

Data mentah yang telah diambil akan melalui beberapa tahap preprocessing untuk mempersiapkan data sebelum klasifikasi yaitu text cleaning, tokenisasi, normalisasi kata gaul, menghapus stopwords, dan stemming

3 Klasifikasi Sentimen dan Hasil Akhir

Pada tahap ini, data yang sudah dipreprocessing akan melalui proses klasifikasi lengkap yang mencakup:

- Ekstraksi Fitur dengan TF-IDF
- Klasifikasi dengan 2 Model Machine Learning
- Visualisasi dan Download Hasil: Hasil klasifikasi ditampilkan dalam bentuk tabel untuk 10 Teratas dan pie chart sebagai visualisasi hasil untuk kedua model. Pengguna bisa mengunduh hasil klasifikasi yang lengkap dalam format CSV.

Halaman About meliputi penjelasan singkat tentang sistem serta panduan singkat mengenai masing-masing langkah yang akan dilakukan untuk menyelesaikan proses klasifikasi sentimen

Halaman Input link

Langkah 1: Input Link dan Scraping Data

Panduan Singkat:

1. Tempelkan URL aplikasi Google Play Store pada kolom "Link".
2. Tentukan jumlah ulasan yang ingin diambil , Maksimal 20,000 ulasan.
3. Klik "Scrape Data" untuk mengambil ulasan terbaru.
4. Periksa preview data mentah yang tampil di bawah.
5. Jika sudah sesuai, klik "Lanjut ke Preprocessing".

Link Google Play Store:

Contoh: <https://play.google.com/store/apps/details?id=com.gojek.app>

Gunakan URL aplikasi Play Store dengan details?id=... — contoh: <https://play.google.com/store/apps/details?id=com.gojek.app..>

Jumlah ulasan yang ingin diambil (maks 20.000):

1000

Scrape Data (Ambil Ulasan Terbaru)

Langkah pertama adalah melakukan data scraping dimana user bisa meng-input link dari aplikasi yang akan diambil ulasannya untuk dilakukan klasifikasi. User juga bisa memilih jumlah ulasan yang ingin diambil. Perhatikan bahwa semakin banyak ulasan yang diambil maka prosesnya akan semakin lama. Setelah selesai, sistem akan menunjukkan 5 data teratas (terbaru) dan bisa lanjut ke langkah kedua yaitu preprocessing

Berhasil mengambil 100 ulasan untuk 'com.shopee.id'.

Preview Data Mentah (5 Baris Teratas)

Ulasan Asli	Rating
tempat belanja yg banyak diskon	5
belanja lebih mudah...scrol²..tf..barang datang	5
sangat puas stiap blanja	5
Mantap belanja di shopee	5
mantap	5

Lanjut ke Preprocessing (Langkah 2)

Halaman Pre-processing

Setelah sistem berhasil mengambil data ulasan melalui proses data scraping, sistem akan lanjut dengan melakukan pre-processing terhadap data yang sudah diambil, meliputi langkah-langkah yang disebutkan di halaman about seperti text cleaning, normalisasi kata gaul, dll. Kemudian sistem akan menunjukkan hasilnya di halaman pre-processing ini. Berikut gambar halaman hasil preprocessing:

Langkah 2: Preprocessing		
Panduan Singkat: <ul style="list-style-type: none">• Data mentah dibersihkan melalui: Text cleaning, Tokenisasi, Normalisasi kata gaul, Stopword Removal, dan Stemming.• Bandingkan kolom "Ulasan Asli" dengan "Teks Akhir" untuk melihat hasil pembersihan.• Jika data tidak muncul, kembali ke halaman Scraping untuk mengambil ulang data.• Jika sudah sesuai, lanjutkan ke proses Klasifikasi (Langkah 3).		
Perbandingan Data (10 Baris Teratas)		
Ulasan Asli (Mentah)	Rating	Teks bersih
mantab	5	mantab
pengalaman belanja sebagian besar memuaskan, semoga terus ditingkatkan layanannya.	5	alam belanja muas moga tingkat layan
pesanan sudah sampai, cukup puas	5	pesan puas
ok	5	ok
minta perbarui terus, tidak bisa untuk belanja. sistem pembayarannya sering error	1	baru belanja sistem bayar error
Good penjualan online	5	good jual online
SANAGAT PUAS DAN TERPERCAYA	5	sanagat puas percaya
ok shoppe...	5	ok shoppe
bosen bgt sm pngirimn yg lmbat spx anteraja ap lh itu pokony skrng shopee lmbat pngirimn gk ky dlu 3hr smpe	1	bosan banget pngirimn lmbat spx anteraja pokony shopee lmbat pngirimn kaya
aku suka apk ini murah sesuai pesanan ada potongan harga blnja lebih hemat love you deh buat apk shopi	5	suka apk murah sesuai pesan potong harga belanja hemat love you deh apk shopi
Lanjut ke Klasifikasi Sentimen (Langkah 3)		

Setelah selesai, sistem akan menunjukkan 10 ulasan terbaru dan bisa mengklik tombol lanjut untuk melakukan proses klasifikasi sentimen

Halaman Hasil Akhir

Langkah 3: Hasil Klasifikasi Sentimen

Panduan Singkat:

- Dua model digunakan: XGBoost dan Adaboost.
- Lihat "Distribusi Sentimen" untuk persentase Positif vs Negatif per model.
- Tabel "Preview Hasil" menampilkan prediksi per ulasan untuk kedua model.
- Gunakan tombol "Download Hasil Lengkap (CSV)" untuk mengunduh semua prediksi.

Distribusi Sentimen



XGBoost

Positif: 67.0%
Negatif: 33.0%



AdaBoost

Positif: 79.0%
Negatif: 21.0%

[Download Hasil Lengkap \(CSV\)](#)

Preview Hasil Klasifikasi

Menampilkan 10 ulasan teratas dengan prediksi dari dua model.

Ulasan Asli	XGBoost	AdaBoost
mantab	Positif	Positif
pengalaman belanja sebagian besar memuaskan, semoga terus ditingkatkan layanannya.	Positif	Positif
pesanan sudah sampai, cukup puas	Positif	Positif
ok	Positif	Positif
minta perbarui terus, tidak bisa untuk belanja. sistem pembayarannya sering error	Negatif	Negatif
Good penjualan online	Positif	Positif
SANAGAT PUAS DAN TERPERCAYA	Positif	Positif
ok shoppe...	Positif	Positif
bosen bgt sm pngirimn yg lmbat spx anteraja ap lh itu pokony skrng shopee lmbat pngirimn gk ky dlu ...	Positif	Positif
aku suka apk ini murah sesuai pesanan ada potongan harga blnja lebih hemat love you deh buat apk sh...	Negatif	Negatif

Halaman hasil akhir menunjukkan hasil prediksi dari kedua model untuk 10 ulasan teratas, serta menyediakan pie chart dari masing-masing model untuk melihat distribusinya. Untuk melihat hasil secara keseluruhan, user bisa mengklik tombol download csv yang disediakan dibawah hasil pie chart

CSV Hasil Akhir

Setelah download csv, bisa membuka file excel yang sudah didownload. Di dalam file excel meliputi ulasan dan rating yang diambil kemudian proses setiap tahapan preprocessing dan hasil prediksi dari model Adaboost dan XGBoost.

	A	B	C	D	E	F	G	H	I
1	content	score	Teks_Bers	Tokens_Fi	Teks_Bers	text_for_v	Prediksi_X	Prediksi_Adaboost	
2	mantab	5	mantab	['mantab']	mantab	mantab	Positif	Positif	
3	pengalam	5	pengalam	['pengalar	alam belar	alam belar	Positif	Positif	
4	pesanan s	5	pesanan s	['pesanan	pesan pua	pesan pua	Positif	Positif	
5	ok	5	ok	['ok']	ok	ok	Positif	Positif	
6	minta perl	1	minta perl	['perbaru	baru belar	baru belar	Negatif	Negatif	
7	Good penj	5	good penj	['good', 'p	good jual	good jual	Positif	Positif	
8	SANAGAT	5	sanagat pi	['sanagat',	sanagat pi	sanagat pi	Positif	Positif	
9	ok shoppe	5	ok shoppe	['ok', 'sho	ok shoppe	ok shoppe	Positif	Positif	
10	bosen bgt	1	bosen bgt	['bosan', 'l	bosan bar	bosan bar	Positif	Positif	
11	aku suka a	5	aku suka a	['suka', 'a	suka apk r	suka apk r	Negatif	Negatif	
12	mitra bisn	5	mitra bisn	['mitra', 'b	mitra bisn	mitra bisn	Negatif	Positif	
13	bagus	5	bagus	['bagus']	bagus	bagus	Positif	Positif	
14	selalu sesu	5	selalu sesu	['sesuai', 'l	sesuai bar	sesuai bar	Positif	Negatif	
15	baik ji	5	baik ji	['ji']	ji	ji	Positif	Positif	
16	terlalu lar	5	terlalu lar	['shoppe',	shoppe pi	shoppe pi	Negatif	Negatif	
17	biaya laya	1	biaya laya	['biaya', 'l	biaya laya	biaya laya	Positif	Positif	
18	mudah me	5	mudah me	['mudah',	mudah ba	mudah ba	Positif	Positif	
19	shopee lay	1	shopee lay	['shopee',	shopee lay	shopee lay	Negatif	Negatif	
20	sangat me	5	sangat me	['memuda	mudah se	mudah se	Positif	Positif	