PERBANDINGAN ANALISIS SENTIMEN TERHADAP ULASAN PELANGGAN AMAZON ANTARA ALGORITMA LOGISTIC REGRESSION DAN MULTINOMIAL NAIVE BAYES

SKRIPSI

Diajukan sebagai salah satu syarat untuk memperoleh
Gelar Sarjana Komputer (S.Kom.)

Sarah Nabila Khairunisa
00000042223

PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG
2024

PERBANDINGAN ANALISIS SENTIMEN TERHADAP ULASAN PELANGGAN AMAZON ANTARA ALGORITMA LOGISTIC REGRESSION DAN MULTINOMIAL NAIVE BAYES

SKRIPSI

Diajukan sebagai salah satu syarat untuk memperoleh
Gelar Sarjana Komputer (S.Kom.)

Sarah Nabila Khairunisa
00000042223

PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG
2024

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama	:	Sarah Nabila Khairunisa
Nomor Induk Mahasiswa	:	00000042223
Program Studi	:	Informatika

Skripsi dengan judul:

Perbandingan Analisis Sentimen terhadap Ulasan Pelanggan Amazon antara Algoritma Logistic Regression dan Multinomial Naive Bayes

merupakan hasil karya saya sendiri bukan plagiat dari karya ilmiah yang ditulis oleh orang lain, dan semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar serta dicantumkan di Daftar Pustaka.

Jika di kemudian hari terbukti ditemukan kecurangan/ penyimpangan, baik dalam pelaksanaan Skripsi maupun dalam penulisan laporan Skripsi, saya bersedia menerima konsekuensi dinyatakan TIDAK LULUS untuk Tugas akhir yang telah saya tempuh.

Tangerang, 3 April 2024

(Sarah Nabila Khairunisa)

HALAMAN PENGESAHAN

Skripsi dengan judul

PERBANDINGAN ANALISIS SENTIMEN TERHADAP ULASAN PELANGGAN AMAZON ANTARA ALGORITMA LOGISTIC REGRESSION DAN MULTINOMIAL NAIVE BAYES

oleh

Nama	:	Sarah Nabila Khairunisa
NIM	:	00000042223
Program Studi	:	Informatika
Fakultas	:	Fakultas Teknik dan Informatika

Telah diujikan pada hari Selasa, 4 Juni 2024
Pukul 15.00 s/d 17.00 dan dinyatakan
LULUS
Dengan susunan penguji sebagai berikut

Ketua Sidang

(Dr. Ir. Winarno, M.Kom.)
NIDN: 330106002

Penguji

(Sy. Yuliani Yakub, S.Kom., M.T., Ph.D.)
NIDN: 0411037904

Pembimbing

(Yaman Khaeruzzaman, M.Sc.)
NIDN: 0413057104

P/t. Ketua Program Studi Informatika,

(Dr. Eng. Niki Prastomo, S.T., M.Sc.)
NIDN: 0419128203

HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH

Yang bertanda tangan di bawah ini:

Nama	:	Sarah Nabila Khairunisa
NIM	:	00000042223
Program Studi	:	Informatika
Jenjang	:	S1
Jenis Karya	:	Skripsi

Menyatakan dengan sesungguhnya bahwa:

☑ Saya bersedia memberikan izin sepenuhnya kepada Universitas Multimedia Nusantara untuk mempublikasikan hasil karya ilmiah saya di repositori Knowledge Center, sehingga dapat diakses oleh Civitas Akademika/Publik. Saya menyatakan bahwa karya ilmiah yang saya buat tidak mengandung data yang bersifat konfidensial dan saya juga tidak akan mencabut kembali izin yang telah saya berikan dengan alasan apapun.

☐ Saya tidak bersedia karena dalam proses pengajuan untuk diterbitkan ke jurnal/konferensi nasional/internasional (dibuktikan dengan *letter of acceptance*)**.

Tangerang, 3 April 2024

Yang menyatakan

Sarah Nabila Khairunisa

** Jika tidak bisa membuktikan LoA jurnal/HKI selama enam bulan ke depan, saya bersedia mengizinkan penuh karya ilmiah saya untuk diunggah ke KC UMN dan menjadi hak institusi UMN.

Halaman Persembahan / Motto

”The best way to not being hopeless is to get up and do something.”

Barrack Obama

KATA PENGANTAR

Puji Syukur atas berkat dan rahmat kepada Tuhan Yang Maha Esa, atas selesainya penulisan laporan Skripsi ini dengan judul: Perbandingan Analisis Sentimen terhadap Ulasan Pelanggan Amazon antara Algoritma Logistic Regression dan Multinomial Naive Bayes dilakukan untuk memenuhi salah satu syarat untuk mencapai gelar Sarjana Komputer Jurusan Informatika Pada Fakultas Teknik dan Informatika Universitas Multimedia Nusantara. Saya menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan skripsi ini, sangatlah sulit bagi saya untuk menyelesaikan skripsi ini. Oleh karena itu, saya mengucapkan terima kasih kepada:

1. Bapak Dr. Ninok Leksono, selaku Rektor Universitas Multimedia Nusantara.
2. Bapak Dr. Eng. Niki Prastomo, S.T., M.Sc., selaku Dekan Fakultas Teknik dan Informatika dan Ketua Program Studi Informatika Universitas Multimedia Nusantara.
3. Bapak Yaman Khaeruzzaman, M.Sc. sebagai dosen pembimbing saya yang telah memberikan bimbingan untuk menyelesaikan skripsi saya.
4. Keluarga dan teman dekat saya yang telah memberikan semangat sehingga saya dapat menyelesaikan skripsi ini.

Semoga skripsi ini bermanfaat, baik sebagai sumber informasi maupun sumber inspirasi, bagi para pembaca.

Tangerang, 3 April 2024

Sarah Nabila Khairunisa

PERBANDINGAN ANALILSIS SENTIMEN TERHADAP ULASAN PELANGGAN AMAZON ANTARA ALGORITMA LOGISTIC REGRESSION DAN MULTINOMIAL NAIVE BAYES

Sarah Nabila Khairunisa

ABSTRAK

Amazon adalah salah satu penyedia jasa penjualan daring terbersar di dunia. Amazon memiliki banyak pembeli dan berbagai macam produk dan layanan yang dijual. Terkadang para pembeli amazon bingung saat ingin membeli produk yang cocok untuk mereka. Oleh karena itu pembeli harus bisa mengetahui ulasan pelanggan sebelumnya sebelum membeli produk. Pembeli harus bisa membedakan ulasan positif dan negatif agar bisa mengetahui apakah produk yang ingin dibeli bagus atau tidak. Pada penelitian ini, dibuat model analisis sentimen untuk mengetahui apakah sebuah ulasan bersifat positif atau negatif dengan membandingkan 2 algoritma yaitu *Logistic Regression* dan *Multinomial Naive Bayes*. Skenario pengujian dilakukan dengan rasio data latih dan data uji sebesar 80:20 dan 70:30 untuk kedua algoritma. Setelah itu, pada masing-masing model diterapkan metode *Synthetic Minority Oversampling Technique* (SMOTE) untuk menyeimbangkan antara data-data dengan kelas positif dan negatif pada kumpulan data latih. Hasil klasifikasi menunjukkan bahwa model dengan algoritma *Logistic Regression* dengan rasio data latih dan data uji sebesar 80:20 dan telah diterapkan SMOTE memiliki hasil akurasi yang paling tinggi yaitu sebesar 86.47%. Hasil presisi pada model tersebut adalah 56.31%, *recall* sebesar 72.96% dan *f1-score* sebesar 63.56%.

Kata kunci: Amazon, Analisis Sentimen, *Logistic Regression*, *Multinomial Naive Bayes*, Python, *Synthetic Minority Oversampling Technique*

THE COMPARISON BETWEEN LOGISTIC REGRESSION AND MULTINOMIAL NAIVE BAYES IN THE SENTIMENT ANALYSIS OF AMAZON CUSTOMER REVIEWS

Sarah Nabila Khairunisa

ABSTRACT

Amazon is one of the biggest e-commerce platforms in the world. It has a lot of customers and offers a lot of products and services for sale. Sometimes Amazon customers get confused while picking the right product for them. Therefore, customers have to know previous customer reviews before buying a product. Customers should be able to tell the difference between positive and negative reviews to determine if the product they want to buy is good or bad. In this research, a sentiment analysis model was made to determine if a review is positive or negative by comparing 2 different algorithms which are Logistic Regression and Multinomial Naive Bayes. There were a few different scenarios, first with 80:20 train-test split and second with 70:30 train-test split for both algorithms. Then, for each model, *Synthetic Minority Oversampling Technique* (SMOTE) was implemented to balance the amount of positive-labelled reviews and negative-labelled reviews in the train data. The classification result showed that the model with Logistic Regression and 80:20 train-test data ratio that had been implemented SMOTE had the biggest accuracy which was 86.47%. The precision was 56.31%, the recall was 72.96% and the f1-score was 63.56%.

Keywords: Amazon, *Logistic Regression*, *Multinomial Naive Bayes*, Python, *Synthetic Minority Oversampling Technique*

DAFTAR ISI

DAFTAR GAMBAR

DAFTAR TABEL

DAFTAR KODE

DAFTAR LAMPIRAN