ANALISIS SENTIMEN PENGGUNA TWITTER TERHADAP KEBOCORAN DATA REGISTRASI SIM MENGGUNAKAN METODE NAÏVE BAYES

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana

Program Studi S1 Informatika

diajukan oleh

RAIHAN ALFAIN SHUBHIY

18.11.2333

Kepada

FAKULTAS ILMU KOMPUTER

UNIVERSITAS AMIKOM YOGYAKARTA

YOGYAKARTA

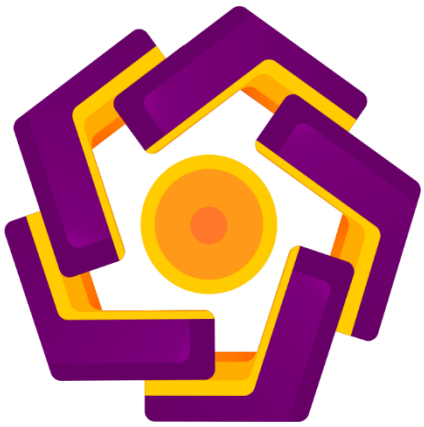
2023

ANALISIS SENTIMEN PENGGUNA TWITTER TERHADAP KEBOCORAN DATA REGISTRASI SIM MENGGUNAKAN METODE NAÏVE BAYES

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana

Program Studi S1 Informatika



diajukan oleh

RAIHAN ALFAIN SHUBHIY

18.11.2333

Kepada

FAKULTAS ILMU KOMPUTER

UNIVERSITAS AMIKOM YOGYAKARTA

YOGYAKARTA

2023

HALAMAN PERSETUJUAN

SKRIPSI

ANALISIS SENTIMEN PENGGUNA TWITTER TERHADAP KEBOCORAN DATA REGISTRASI SIM MENGGUNAKAN METODE NAÏVE BAYES

yang disusun dan diajukan oleh

RAIHAN ALFAIN SHUBHIY

18.11.2333

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 4 April 2023

Dosen Pembimbing,

Theopilus Bayu Sasongko, S.Kom, M.Eng.
NIK. 190302375

HALAMAN PENGESAHAN

SKRIPSI

ANALISIS SENTIMEN PENGGUNA TWITTER TERHADAP KEBOCORAN DATA REGISTRASI SIM MENGGUNAKAN METODE NAÏVE BAYES

yang disusun dan diajukan oleh

RAIHAN ALFAIN SHUBHIY

18.11.2333

Telah dipertahankan di depan Dewan Penguji
pada tanggal 4 April 2023

Susunan Dewan Penguji

Nama Penguji	Tanda Tangan
Arif Dwi Laksito, M.Kom. **NIK. 190302150**	
Supriatin, M.Kom. **NIK. 190302239**	
Theopilus Bayu Sasongko, S.Kom., M.Eng. **NIK. 190302375**	

Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 4 April 2023

DEKAN FAKULTAS ILMU KOMPUTER

Hanif Al Fatta,S.Kom., M.Kom.
NIK. 190302096

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Raihan Alfain Shubhiy
NIM : 18.11.2333

Menyatakan bahwa Skripsi dengan judul berikut:

Analisis Sentimen Pengguna Twitter Terhadap Kebocoran Data Registrasi SIM Menggunakan Metode Naïve Bayes

Dosen Pembimbing : Theopilus Bayu Sasongko, S.Kom, M.Eng.

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 4 April 2023

Yang Menyatakan,

Raihan Alfain Shubhiy

HALAMAN PERSEMBAHAN

Alhamdulillah, skripsi ini telah saya selesaikan dengan baik dan maksimal. Hal ini tentunya tak lepas dari karunia, rahmat serta hidayah yang telah diberikan oleh Allah swt. Sehingga saya mendapat kemudahan dan kelancaran dalam mengerjakan skripsi ini. Selain itu ada orang-orang hebat yang selalu memberikan dukungan, motivasi, dan bantuan baik secara langsung maupun tidak langsung, antar lain.

1. Orang tua serta keluarga yang selalu memberikan dukungan, motivasi dan doa restu tanpa henti.
2. Bapak Theopilus Bayu Sasongko, S.Kom, M.Eng yang telah membimbing saya dalam peroses pengerjaan skripsi ini.
3. Teman-Teman dari Himpunan Mahasiswa Informatika yang membantu dikala saya mendapati kesulitan dalam pengerjaan skripsi.
4. Shabrina Azzahra, Amd. Keb yang selalu menemani dan membantu saya dalam proses pengerjaan skripsi hingga selesai.
5. Teman-Teman yang tidak bisa saya sebutkan namanya satu-persatu yang telah membantu dan men-*support* baik secara langsung maupun tidak langsung.

KATA PENGANTAR

Puji syukur kehadirat Allah swt. yang telah memberikan karunia, rahmat serta hidayah-nya sehingga penulis dapat menyelesaikan skripsi ini dengan baik dan maksimal. Skripsi yang berjudul **"Analisis Sentimen Pengguna Twitter Terhadap Kebocoran Data Registrasi SIM Menggunakan Metode Naïve Bayes"** ini disusun sebagai salah satu syarat dalam menyelesaikan masa studi program sarjana di Universitas AMIKOM Yogyakarta.

Dalam penulisan skripsi ini, penulis mengucapkan rasa terimakasih atas motivasi, bimbingan, saran dan masukan dari berbagai pihak secara moral maupun spiritual. Pada kesempatan kali ini penulis mengucapkan terimakasih kepada:

1. Bapak Prof. Dr. M. Suyanto, M.M selaku Rektor Universitas AMIKOM Yogyakarta.
2. Bapak Hanif Al Fatta, M.Kom selaku Dekan Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.
3. Ibu Windha Mega Pradya D., M.Kom selaku Ketua Program Studi S1 Informatika Universitas Amikom Yogyakarta.
4. Bapak Theopilus Bayu Sasongko, S.Kom, M.Eng selaku Dosen Pembimbing yang selalu memberikan saran dan masukan dalam proses penulisan skripsi.
5. Bapak Arif Dwi Laksito, M.Kom dan Ibu Supriatin, M.Kom selaku Dosen Penguji yang telah memberikan evaluasi dan saran agar penelitian ini menjadi lebih baik.

Akhir kata, semoga penyusunan skripsi ini dapat bermanfaat bagi pembaca dalam menambah wawasan dan pengetahuan khususnya dalam bodang informatika.

Yogyakarta, 4 April 2023

Penulis

DAFTAR ISI

DAFTAR TABEL

DAFTAR GAMBAR

INTISARI

Perkembangan teknologi digital di Indonesia yang semakin masif membuat masyarakat tidak bisa menolak dampak dari digitalisasi yang dapat membawa ancaman besar seperti ancaman keamanan data pribadi. Pada bulan September 2022, terdapat insiden kebocoran data registrasi kartu sim berisikan data pribadi pengguna yang disebarkan melalui situs gelap. Sebagai salah satu media sosial paling populer di Indonesia, Twitter dijadikan tempat oleh masyarakat untuk menyuarakan opininya terkait isu kebocoran data registrasi sim. Penelitian ini bertujuan untuk menganalisis sentimen dan sebaran kata dari opini pengguna Twitter terkait dengan isu tersebut. Analisis sentimen dilakukan menggunakan pendekatan *machine learning* dengan metode klasifikasi *Naïve Bayes*. Penelitian ini menggunakan 901 data *tweet* yang sudah diberi label. Terdapat ketidakseimbangan antar kelas sentimen pada *dataset* yang digunakan dimana data sentimen positif memiliki jumlah yang jauh lebih sedikit dibanding dengan sentimen negatif dan netral. Sehingga dilakukan teknik *oversampling* menggunakan SMOTE pada data latih untuk membantu algoritma dalam membentuk model klasifikasi. Hasil dari model yang dibangun menggunakan algoritma *Naïve Bayes* mendapatkan nilai *accuracy* sebesar 71%, *precision* 62%, *recall* 74%, dan *f1-score* sebesar 62%. Ketidakseimbangan kelas sentimen membuat algoritma *Naïve Bayes* memiliki performa yang rendah.

Kata kunci: *Naïve Bayes*, SMOTE, Kebocoran Data SIM, Analisis Sentimen, Twitter.

ABSTRACT

The increasingly massive development of digital technology in Indonesia has made people unable to resist the impact of digitalization which can bring great threats such as threats to personal data security. In September 2022, there was an incident of sim card registration data leak containing users' personal data that was spread through dark sites. As one of the most popular social media in Indonesia, Twitter is used by the public to voice their opinions regarding the issue of sim registration data leakage. This study aims to analyze the sentiment and word distribution of Twitter users' opinions related to the issue. Sentiment analysis is conducted using a machine learning approach with the Naïve Bayes classification method. This research uses 901 tweet data that has been labeled. There is an imbalance between sentiment classes in the dataset used where positive sentiment data has a much smaller amount than negative and neutral sentiment. So an oversampling technique using SMOTE is carried out on the training data to help the algorithm in forming a classification model. The results of the model built using the Naïve Bayes algorithm get an accuracy value of 71%, precision 62%, recall 74%, and f1-score of 62%. The imbalance of sentiment classes makes the Naïve Bayes algorithm have low performance.

Keyword: *Naïve Bayes, SMOTE, SIM data leak, sentiment analysis, Twitter.*