ANALISIS SENTIMEN PADA MEDIA SOSIAL TWITTER TERHADAP CYBERBULLYING MENGGUNAKAN ALGORITMA BIDIRECTIONAL LONG SHORT TERM MEMORY

JALUR SCIENTIST

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana

Program Studi Informatika

disusun oleh

ANISA IKA SAFITRI

20.11.3762

Kepada

FAKULTAS ILMU KOMPUTER

UNIVERSITAS AMIKOM YOGYAKARTA

YOGYAKARTA

2024

ANALISIS SENTIMEN PADA MEDIA SOSIAL TWITTER TERHADAP CYBERBULLYING MENGGUNAKAN ALGORITMA BIDIRECTIONAL LONG SHORT TERM MEMORY

JALUR SCIENTIST

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana

Program Studi Informatika

disusun oleh

ANISA IKA SAFITRI

20.11.3762

Kepada

FAKULTAS ILMU KOMPUTER

UNIVERSITAS AMIKOM YOGYAKARTA

YOGYAKARTA

2024

HALAMAN PERSETUJUAN

JALUR SCIENTIST

ANALISIS SENTIMEN PADA MEDIA SOSIAL TWITTER TERHADAP CYBERBULLYING MENGGUNAKAN ALGORITMA BIDIRECTIONAL LONG SHORT TERM MEMORY

yang disusun dan diajukan oleh

ANISA IKA SAFITRI

20.11.3762

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 21 Mei 2024

Dosen Pembimbing,

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

HALAMAN PENGESAHAN

SKRIPSI

ANALISIS SENTIMEN PADA MEDIA SOSIAL TWITTER TERHADAP CYBERBULLYING MENGGUNAKAN ALGORITMA BILSTM

yang disusun dan diajukan oleh

ANISA IKA SAFITRI

20.11.3762

Telah dipertahankan di depan Dewan Penguji
pada tanggal 21 Mei 2024

Susunan Dewan Penguji

Nama Penguji	Tanda Tangan
Nuri Cahyono, M.Kom **NIK. 190302278**	
Subektiningsih, M.Kom **NIK. 190302413**	
Anna Baita, M. Kom **NIK. 190302290**	

Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 21 Mei 2024

DEKAN FAKULTAS ILMU KOMPUTER

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

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Anisa Ika Safitri
NIM : 20.11.3762

Menyatakan bahwa Skripsi dengan judul berikut:

ANALISIS SENTIMEN PADA MEDIA SOSIAL TWITTER TERHADAP CYBERBULLYING MENGGUNAKAN ALGORITMA BIDIRECTIONAL LONG SHORT TERM MEMORY

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, 21 Mei 2024

Yang Menyatakan,

Anisa Ika Safitri

HALAMAN PERSEMBAHAN

Dengan penuh rasa hormat dan syukur kepada Tuhan Yang Maha Esa, saya ingin mengucapkan penghargaan yang tak terhingga kepada semua pihak yang telah memberikan dukungan dan bantuan selama proses penyusunan skripsi ini :

1. Allah SWT yang telah memberikan rahmat, hidayah, petunjuk dan kekuatan-Nya yang tidak terbatas kepada penulis sehingga dapat menyelesaikan skipsi.
2. Terima kasih kepada kedua orang tua yang selalu memberikan doa, dukungan, dan motivasi tiada henti dalam setiap langkah perjalanan hidup saya.
3. Dosen pembimbing Bapak Theopilus Bayu Sasongko, S.Kom., M.Eng. yang telah memberikan arahan, masukan, dan bimbingan yang sangat berharga dalam menuntun saya dalam penelitian ini.
4. Sahabat saya Meita Dwi Rahayu, Avida Dwi Rosanti, Siska Puspita Sari yang selalu memberi dukungan dan semangat.
5. Diri sendiri yang sudah bertahan dan berjuang dalam menyelesaikan skripsi hingga akhir.

KATA PENGANTAR

Puji syukur penulis panjatkan kehadirat Allah SWT, yang telah melimpahkan rahmat dan hidayah-Nya kepada penulis sehingga dapat menyelesaikan penelitian dengan judul "Analisis Sentimen Pada Media Sosial Twitter Terhadap Cyberbullying Menggunakan Algoritma BILSTM" sehingga penulis dapat menyelesaikan penelitian dengan baik. Skripsi ini disusun untuk memenuhi syarat akademis dalam menyelesaikan Program Studi Informatika Strata Satu (S1) pada Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.

Dalam proses penyusunan skripsi ini, Penulis ingin mengucapkan terima kasih atas bantuan, motivasi dan dukungan dari berbagai pihak, antara lain:

1. Bapak Prof. Dr. M. Suyanto,M.M, selaku Rektor Universitas AMIKOM Yogyakarta.
2. Bapak Hanif Al Fatta, S. Kom., M. Kom, selaku Dekan Fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta.
3. Ibu Windha Mega Pradnya Dhuhita, M.Kom. selaku ketua program studi S1 Informatika.
4. Bapak Theopilus Bayu Sasongko, S.Kom.,M.Eng. selaku dosen pembimbing yang telah memberikan arahan, masukan, dan bimbingan kepada penulis dalam menyelesaikan skripsi.
5. Bapak Tristanto Ari Aji, M.Kom selaku Dosen Wali selama menjadi mahasiswa di Amikom.
6. Teman-teman seperjuangan dari kelas 20-IF-08 yang selalu menemani dan memberi support selama perkuliahan.

Penulis menyadari bahwa skripsi ini belum mencapai tingkat kesempurnaan yang diinginkan dan masih memiliki kelemahan yang perlu diperbaiki. Oleh karena itu, penulis mengharapkan kritik dan saran dari pembaca. Semoga hasil penelitian ini memberikan manfaat yang berarti bagi para pembaca.

Yogyakarta, 21 Mei 2024
Penulis

DAFTAR ISI

HALAMAN JUDUL i
HALAMAN PERSETUJUAN ii
HALAMAN PENGESAHAN iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI iv
HALAMAN PERSEMBAHAN v
KATA PENGANTAR vi
DAFTAR ISI vii
DAFTAR TABEL ix
DAFTAR GAMBAR x
DAFTAR LAMPIRAN xi
DAFTAR LAMBANG DAN SINGKATAN xii
INTISARI xiii
ABSTRACT xiv
BAB I PENDAHULUAN 1
1.1 Latar Belakang 1
BAB II LANDASAN TEORI 3
BAB III METODE PENELITIAN 5
3.1 *Dataset* 6
3.2 *Preprocessing* 7
3.3 *Data Balancing* 7
3.4 *Word Embedding* 8
3.5 *Bidirectional Long Short-Term Memory* 8

3.6 *Evaluasion* 8
BAB IV HASIL DAN PEMBAHASAN 10
4.1 *Dataset* 10
4.2 *Preprocessing* 11
4.2.1 *Text Cleaning* 11
4.2.2 *Case folding* 12
4.2.3 *Lemmatization* 12
4.2.4 *Stopword Removal* 12
4.2.5 *Tokenization* 12
4.3 *Word Embedding* 13
4.4 *Data Balancing* 14
4.5 *Modeling(Bidirectional Long Short-Term Memory* (BiLSTM)) 14
4.6 *Evaluation* 15
BAB V PENUTUP 17
5.1 Kesimpulan 17
5.2 Saran 17
REFERENSI 18
LAMPIRAN BUKTI PENDUKUNG 20
a. Letter of Acceptance (LOA) 20
b. Lembar Review dari Penyelenggara Jurnal 20
1. *Reviewer* A 21
2. *Reviewer* B 21
c. Bukti Bahwa Naskah Telah terbit dan terindex 22

DAFTAR TABEL

Tabel 3. 1 Informasi Dataset 6
Tabel 4. 1 Tweet Sebelum dan Sesudah Preprocessing 11
Tabel 4. 2 Tweet Sebelum Preprocessing 10
Tabel 4. 3 Tweet Setelah Preprocessing 11
Tabel 4. 4 Jumlah Data Setiap Kelas 14
Tabel 4. 5 Parameter 15
Tabel 4. 6 Classification Bi-LSTM 16
Tabel 4. 7 Evaluation Metrics 16

DAFTAR GAMBAR

Gambar 3. 1 Alur Penelitian 5
Gambar 3. 2 Visualisasi Sentiment 6
Gambar 3. 3 Confusion Matrix 9
Gambar 4. 1 Wordcloud Sebelum Processing 13
Gambar 4. 2 Wordcloud Setelah Preprocessing 13
Gambar 4. 3 Confusion Matrix 15

DAFTAR LAMPIRAN

Lampiran 1. LOA dari JUTIF 20
Lampiran 2. Review dari Reviewer A 21
Lampiran 3. Review dari Reviewer B 21
Lampiran 4. Artikel terbit di JUTIF 22
Lampiran 5. Akreditasi Jurnal Teknik Informatika (JUTIF) 22
Lampiran 6. Artikel Terbit di Google Scholar 23

DAFTAR LAMBANG DAN SINGKATAN

BiLSTM	Bidirectional Long Short-Term Memory
NLTK	Natural Language Toolkit
TP	True positive
FP	False positive
FN	False negative
TN	True negative

INTISARI

Cyberbullying di media sosial seperti Twitter menjadi masalah sosial yang semakin meningkat di masyarakat saat ini. Cyberbullying memiliki pengaruh negatif pada kesehatan mental, meningkatkan risiko kecemasan, kesedihan, dan bahkan bunuh diri. Tujuan dari penelitian ini adalah untuk mengembangkan model untuk mengklasifikasikan tweet yang mengandung atau tidak mengandung cyberbullying dengan menerapkan teknik BiLSTM pada analisis sentimen di Twitter. Dalam penelitian ini Word2Vec digunakan untuk pembobotan setiap kata pada sebuah tweet. Tahap awal pada penelitian ini dilakukan adalah pengumpulan data dengan jumlah dataset sebanyak 47.692 tweet yang dihasilkan oleh Kaggle, preprocessing yang terdiri dari data cleaning, remove duplicate, case folding, tokenizing, stopword removal dan lemmatization, classification dan evaluasi. Penelitian ini menggunakan metode Bidirectional Long Short-Term Memory (Bi-LSTM) dan mengidentifikasi pola-pola yang terkait dengan tindakan bullying di media sosial. Pengujian menggunakan Confusion Matrix dan hasil pada klasifikasi menunjukkan accuracy sebesar 82,29%, presisi 82,04%, recall 81,95% dan F1-Score 81,89%. Teknik analisis sentimen ini diharapkan dapat menjadi langkah awal untuk memerangi dan menghindari cyberbullying di platform Twitter. Dari beberapa pengujian algoritma referensi yang ada, akurasi klasifikasi yang dilakukan termasuk memiliki performa yang bagus.

Kata kunci: Bi-lstm, Cyberbullying, Analisis Sentimen, Twitter, Word2vec.

ABSTRACT

Cyberbullying on social media such as Twitter is becoming an increasing social problem in today's society. Cyberbullying has a negative influence on mental health, increasing the risk of anxiety, sadness, and even suicide. The purpose of this research is to develop a model to classify tweets that contain or do not contain cyberbullying by applying the BiLSTM technique to sentiment analysis on Twitter. In this research, Word2Vec is used to weight each word in a tweet. The initial stage in this research is data collection with a total dataset of 47,692 tweets generated by Kaggle, preprocessing which consists of data cleaning, removing duplicates, case folding, tokenizing, stopword removal and lemmatization, classification and evaluation. This research uses the Bidirectional Long Short-Term Memory (Bi-LSTM) method and identifies patterns associated with bullying on social media. Testing uses Confusion Matrix and the results on classification show accuracy of 82.29%, precision of 82,04%, recall of 81,95% and F1-Score 81,89%. This sentiment analysis technique is expected to be the first step to combat and avoid cyberbullying on the Twitter platform. From several tests of existing reference algorithms, the classification accuracy performed includes having good performance.

Keyword: Bi-Lstm, Cyberbullying, Sentiment Analysis, Twitter, Word2vec.