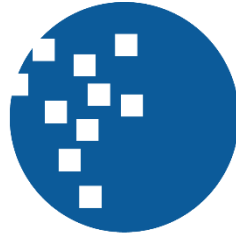


**DETEKSI KEKERASAN SEKSUAL VERBAL BERBAHASA
INDONESIA DARI MEDIA SOSIAL TWITTER
MENGUNAKAN AWD-LSTM**



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

SKRIPSI

Alvian Pratama

0000029687

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG**

2022

**DETEKSI KEKERASAN SEKSUAL VERBAL BERBAHASA
INDONESIA DARI MEDIA SOSIAL TWITTER
MENGUNAKAN AWD-LSTM**



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

SKRIPSI

Diajukan sebagai Salah Satu Syarat untuk Memperoleh

Gelar Sarjana Komputer (S.Kom.)

Alvian Pratama

0000029687

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA**

TANGERANG

2022

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Alvian Pratama

Nomor Induk Mahasiswa : 0000029687

Program studi : Sistem Informasi

Skripsi dengan judul: Deteksi Kekerasan Seksual Verbal Berbahasa Indonesia Dari Media Sosial Twitter Menggunakan AWD-LSTM 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, 15 Juni 2022



(Alvian Pratama)

HALAMAN PENGESAHAN

Skripsi dengan judul Deteksi Kekerasan Seksual Verbal Berbahasa Indonesia Dari Media Sosial Twitter Menggunakan AWD-LSTM

Oleh
Nama : Alvian Pratama
NIM : 00000029687
Program Studi : Sistem Informasi
Fakultas : Teknik dan Informatika

Telah diujikan pada hari Selasa, 28 Juni 2022
Pukul 10.00 s.d 12.00 dan dinyatakan
LULUS
Dengan susunan pengujian sebagai berikut.

Ketua Sidang

Penguji



11/07/2022

Johan Setiawan, S.Kom., M.M., M.B.A.

11-7-2022

Raymond Sunardi Oetama, MCIS.

Pembimbing



12/07/2022

Monika Evelin Johan, S.Kom., M.M.S.I.

Ketua Program Studi Sistem Informasi

UNIVERSITAS
MULTIMEDIA
NUSANTARA



12/07/2022

Ririn Ikana Desanti, S.Kom., M.Kom.

HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Sebagai civitas academica Universitas Multimedia Nusantara, saya yang bertanda tangan di bawah ini:

Nama : Alvian Pratama
NIM : 00000029687
Program Studi : Sistem Informasi
Fakultas : Teknik & Informatika
Jenis Karya : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Multimedia Nusantara Hak Bebas Royalti Noneksklusif (*Non-exclusive Royalty-Free Right*) atas karya ilmiah saya yang berjudul. Deteksi Kekerasan Seksual Verbal Berbahasa Indonesia Dari Media Sosial Twitter Menggunakan AWD-LSTM. Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Noneksklusif ini, Universitas Multimedia Nusantara berhak menyimpan, mengalihmediakan/mengalihformatkan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan memublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta. Demikian pernyataan ini saya buat dengan sebenarnya.

Tangerang, 15 Juni 2022

Yang menyatakan,



(Alvian Pratama)

U M M N
U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

KATA PENGANTAR

Puji dan syukur saya ucapkan kepada Tuhan Yang Maha Esa, karena atas dasar Penyertaan-Nya, saya mampu untuk menyelesaikan proposal yang berjudul “Deteksi Kekerasan Seksual Verbal Berbahasa Indonesia Dari Media Sosial Twitter Menggunakan AWD-LSTM” dengan tepat waktu.

Proposal ini diajukan sebagai salah satu syarat untuk memenuhi kewajiban dari mata kuliah yang bernama *Information System Research and Methodology*. Selain itu, proposal ini juga akan dipakai untuk diajukan sebagai bahan untuk skripsi mendatang. Penulisan dari proposal ini dapat diselesaikan dengan bantuan oleh beberapa peran berikut ini :

1. Dr. Ninok Leksono, selaku Rektor Universitas Multimedia Nusantara.
2. Dr. Eng. Niki Prastomo, S.T., M.Sc., selaku Dekan Fakultas Universitas Multimedia Nusantara.
3. Ibu Ririn Ikana Desanti, S.Kom., M.Kom., selaku Ketua Program Studi Universitas Multimedia Nusantara.
4. Ibu Monika Evelin Johan, S.Kom., M.M.S.I., sebagai Pembimbing pertama yang telah memberikan bimbingan, arahan, dan motivasi atas terselesainya tugas akhir ini.
5. Kepada keluarga dan para rekan yang berpartisipasi dalam memberikan semangat dan doa dalam penyelesaian proposal ini.

Penulis berharap bahwa penulisan dari proposal ini dapat memberikan informasi dan inspirasi yang berguna untuk masa mendatang. Mohon maaf apabila masih terdapat kekurangan dalam penulisan.

Tangerang, 15 Juni 2022



Alvian Pratama

DETEKSI KEKERASAN SEKSUAL VERBAL BERBAHASA INDONESIA DARI MEDIA SOSIAL TWITTER MENGUNAKAN AWD-LSTM

(Alvian Pratama)

ABSTRAK

Penetrasi media sosial di Negara Indonesia sangat luas cakupannya. Sebanyak 61,8 persen penduduk Indonesia sudah pernah menyentuh media sosial. Perkembangan dari media sosial yang terjadi di Negara Indonesia begitu cepat sehingga penyebaran informasi tidak lagi dapat dikontrol oleh pemerintah. Oleh karena hal tersebut, terjadi banyak bentuk kekerasan yang terjadi di media sosial. Salah satu kasus yang marak terjadi adalah kasus yang bertemakan kekerasan seksual verbal yang terjadi di media sosial. Diperkirakan pada tahun 2021, terjadi peningkatan kekerasan seksual yang dimana bertambah sebanyak 241 kasus.

Pada umumnya, kasus dari kekerasan seksual akan ditangani oleh lembaga-lembaga pemerintah ataupun lembaga masyarakat yang mengkhususkan terhadap kategori tersebut. Namun dengan semakin banyaknya kasus kekerasan seksual yang terjadi semenjak pandemic *Covid-19*, diperkirakan bahwa lembaga – lembaga tersebut akan kewalahan dalam menangani kasus kekerasan seksual. Oleh karena itu dibutuhkan kecerdasan buatan untuk dapat menangani kasus – kasus dari kekerasan seksual tersebut yang dimana dapat mendeteksi kekerasan seksual yang terjadi di media sosial.

Penelitian ini dibuat dengan tujuan untuk mendeteksi kekerasan seksual yang terdapat dalam pesan di media sosial dengan menggunakan algoritma *deep learning* yang bernama *Average Stochastic Gradient Descent Weight-Dropped LSTM* (AWD-LSTM). Permodelan yang dibuat akan menggunakan dataset yang memiliki konten berupa pesan ataupun komentar yang bersumber dari media sosial Twitter. Setelah permodelan selesai dilakukan, model yang sudah selesai dibangun akan diimplementasikan dalam bentuk web yang dimana dapat menerima inputan berupa konten yang berisikan sebuah teks. Implementasi dari web tersebut akan menerima inputan berupa teks yang kemudian akan diolah untuk dapat mengetahui apakah inputan tersebut mengandung unsur kekerasan seksual.

Kata kunci: AWD-LSTM, Flask, Kekerasan Seksual, Klasifikasi, Media Sosial

**DETECTION OF VERBAL SEXUAL VIOLENCE IN
INDONESIAN LANGUAGE FROM TWITTER SOCIAL
MEDIA USING AWD-LSTM**

(Alvian Pratama)

ABSTRACT (English)

Social media penetration in Indonesia is very broad in a scope. As many as 61.8 percent of Indonesia's population has touched social media. The development of social media that occurs in Indonesia is so fast that the dissemination of information can no longer be controlled by the government. Because of this, there are many forms of violence that occur on social media. One of the most common cases is a case with the theme of sexual violence that occurs on social media. It is estimated that in 2021, there will be an increase in sexual violence which will increase by 241 cases.

In general, cases of sexual violence will be handled by government agencies or community institutions that specialize in this category. However, with the increasing number of cases of sexual violence that have occurred since the Covid-19 pandemic, it is estimated that these institutions will be overwhelmed in handling cases in the form of sexual violence. Therefore, artificial intelligence is needed so it can detect sexual violence that occurs on social media.

This study was made with the aim of detecting sexual violence contained in messages on social media using a deep learning algorithm called Average Stochastic Gradient Descent Weight-Dropped LSTM (AWD-LSTM). The model made will use a dataset that has content in the form of messages or comments sourced from Twitter social media. After the modeling is completed, the completed model will be implemented in the form of a web which can accept input in the form of text content. The implementation of the web will receive input in the form of texts which will then be processed to determine whether the input contains elements of sexual violence.

Keywords: AWD-LSTM, Classification, Flask, Sexual Violence, Social Media

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

DAFTAR ISI

HALAMAN PERNYATAAN TIDAK PLAGIAT	i
HALAMAN PENGESAHAN.....	ii
HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS	iii
KATA PENGANTAR	iv
ABSTRAK	v
<i>ABSTRACT (English)</i>	vi
DAFTAR ISI.....	vii
DAFTAR TABEL.....	ix
DAFTAR GAMBAR	x
DAFTAR RUMUS	xi
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah.....	3
1.4 Tujuan Penelitian.....	4
1.6 Sistematika Penulisan.....	4
BAB II LANDASAN TEORI.....	6
2.1 Tinjauan Teori	6
2.1.1 Media Sosial.....	6
2.1.2 Kekerasan Seksual	6
2.1.3 Kekerasan Berbasis Gender Online	7
2.1.4 Kekerasan Seksual Verbal.....	7
2.1.5 Knowledge Discovery in Databases.....	7
2.1.6 Text Preprocessing	9
2.1.7 Average Stochastic Gradient Descent Weight-Dropped LSTM (AWD-LSTM)	10
2.1.8 Flask.....	11
2.2 Penelitian Terdahulu.....	12
BAB III METODOLOGI PENELITIAN	25
3.1 Gambaran Umum Objek Penelitian	25

3.2	Metode Penelitian.....	25
3.2.1	Alur Penelitian.....	26
3.3	Variabel Penelitian.....	31
3.3.1	Variabel Independen.....	31
3.3.2	Variabel Dependen.....	32
3.4	Teknik Pengumpulan Data.....	32
3.5	Teknik Pengambilan Sampel.....	32
3.6	Teknik Analisis Data.....	33
3.7	Kerangka Teori.....	34
3.7.1	Rumusan Masalah 1.....	34
3.7.2	Rumusan Masalah 2.....	35
BAB IV ANALISIS DAN HASIL PENELITIAN.....		36
4.1	Pengumpulan Data.....	36
4.2	Labelling Data.....	36
4.3	Data Preprocessing.....	37
4.4	Pelatihan Model.....	41
4.5	Evaluasi Model.....	48
4.6	Implementasi Model.....	52
4.7	Hasil Analisis dan Diskusi.....	57
BAB V KESIMPULAN DAN SARAN.....		65
5.1	Kesimpulan.....	65
5.2	Saran.....	66
DAFTAR PUSTAKA.....		67
LAMPIRAN.....		72

DAFTAR TABEL

Tabel 2.1. Penelitian Terdahulu.....	12
Tabel 3.1. Perbandingan Arsitektur Data Mining.....	25
Tabel 3.2. Perbandingan Arsitektur Algoritma.....	33
Tabel 4.1. Perbandingan Sebelum dan Sesudah Preprocessing.....	38
Tabel 4.2. Perbandingan Hyperparameter dari Model.....	45
Tabel 4.3. Tabel Perbandingan Hasil dari Hyperparameter Model.....	51
Tabel 4.4. Confusion Matrix Neural Network.....	52
Tabel 4.5. Hasil Prediksi Menggunakan Sampel Baru.....	55
Tabel 4.5. Hasil Prediksi Menggunakan Sampel Baru.....	55
Tabel 4.6. Tabel Perbandingan Model <i>LSTM</i> dengan <i>AWD-LSTM</i>	57
Tabel 4.7. Tabel Perbandingan Penelitian dengan Penelitian Terdahulu.....	58
Tabel 4.8. Bobot Nilai Jawaban.....	60
Tabel 4.9. Kriteria Nilai yang diberikan.....	60
Tabel 4.10. Isi dari Kuesioner	61
Tabel 4.11. Hasil Kuesioner Untuk Pertanyaan 1.....	62
Tabel 4.12. Hasil Kuesioner Untuk Pertanyaan 2.....	62
Tabel 4.13. Hasil Kuesioner Untuk Pertanyaan 3.....	63

UMMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA

DAFTAR GAMBAR

Gambar 2.1. Arsitektur Knowledge Discovery in Databases.....	7
Gambar 2.2. Arsitektur AWD-LSTM.....	11
Gambar 3.1. Flowchart Implementasi Model.....	31
Gambar 3.1. Kerangka Teori.....	34
Gambar 4.1. Data yang Disimpan dalam Bentuk .CSV.....	36
Gambar 4.2. Data yang Sudah Diberikan Label.....	37
Gambar 4.3. Script Untuk Proses Preprocessing.....	39
Gambar 4.4. Data Sebelum dan Sesudah Preprocessing.....	39
Gambar 4.5. Jumlah Data Sebelum dan Sesudah Preprocessing.....	40
Gambar 4.6. Perbandingan Label Sebelum Proses Preprocessing.....	40
Gambar 4.7. Perbandingan Label Sesudah Proses Preprocessing.....	41
Gambar 4.8. Jumlah Data Train.....	42
Gambar 4.9. Jumlah Data Test.....	42
Gambar 4.10. Perbandingan Label Pada Data Train.....	43
Gambar 4.11. Perbandingan Label pada Data Test.....	43
Gambar 4.12. Script Untuk AWD-LSTM.....	44
Gambar 4.13. Script Untuk Hyperparameter Model.....	45
Gambar 4.14. Grafik Perbandingan Train_Loss dan Valid Loss Hyperparameter 1.....	46
Gambar 4.15. Grafik Perbandingan Train_Loss dan Valid Loss Hyperparameter 2.....	46
Gambar 4.16. Grafik Perbandingan Train_Loss dan Valid Loss Hyperparameter 3.....	47
Gambar 4.17. Grafik Perbandingan Train_Loss dan Valid Loss Hyperparameter 4.....	47
Gambar 4.18. Grafik Perbandingan Train_Loss dan Valid Loss Hyperparameter 5.....	47
Gambar 4.19. Hasil dari Hyperparameter 1.....	48
Gambar 4.20. Hasil dari Hyperparameter 2.....	49
Gambar 4.21. Hasil dari Hyperparameter 3.....	49
Gambar 4.22. Hasil dari Hyperparameter 4.....	50
Gambar 4.23. Hasil dari Hyperparameter 5.....	50
Gambar 4.24. Hasil prediksi teks dari Model.....	52
Gambar 4.25. Form Untuk Pengisian Input Berupa Teks.....	53
Gambar 4.26. Contoh Teks yang Mengandung Kekerasan Seksual Verbal.....	53
Gambar 4.27. Output Dari Teks yang Mengandung Kekerasan Seksual Verbal..	54
Gambar 4.28. Contoh Teks yang Tidak Mengandung Kekerasan Seksual Verbal.....	54
Gambar 4.29. Output Dari Teks yang Tidak Mengandung Kekerasan Seksual Verbal.....	54
Gambar 4.30. Infografis Kekerasan Seksual yang Terjadi Dari Tanggal 29 Maret 2022 sampai 5 April 2022.....	56

DAFTAR RUMUS

Rumus 3.1. Rumus Perhitungan Akurasi	30
Rumus 3.2. Rumus Perhitungan Precision	30
Rumus 3.3. Rumus Perhitungan Recall	30
Rumus 3.4. Rumus Perhitungan F1-Score	30
Rumus 4.1. Rumus Perhitungan Nilai Kriteria.....	61



UMMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA