IMPLEMENTASI MODEL *INDOBERT* PADA ANALISIS SENTIMEN TERHADAP ISU FEMINISME DI TWITTER

SKRIPSI

Diajukan untuk memenuhi salah satu syarat kelulusan
di Program Studi Sains Data

Disusun Oleh:

BRESCIA AYUNDINA YUNIAROSSY

20083010021

UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN" JAWA TIMUR

FAKULTAS ILMU KOMPUTER

PROGRAM STUDI SAINS DATA

SURABAYA

2024

LEMBAR PENGESAHAN

IMPLEMENTASI MODEL *INDOBERT* PADA ANALISIS SENTIMEN TERHADAP ISU FEMINISME DI TWITTER

SKRIPSI

Diajukan untuk memenuhi salah satu syarat memperoleh gelar Sarjana Sains Data
pada : Senin, 15 Juli 2024

Program Studi S-1 Sains Data
Fakultas Ilmu Komputer
Universitas Pembangunan Nasional Veteran Jawa Timur
Surabaya

Oleh :

BRESCIA AYUNDINA YUNIAROSSY

NPM. 20083010021

Disetujui oleh Tim Penguji Skripsi :

Penguji 1	Penguji 2
Dr. Ir. Mohammad Idhom, SP., S.Kom., MT	Trimono, S.Si., M.Si
NIP. 198303102021211006	NIP. 199509082022031003
Pembimbing 1	**Pembimbing 2**
Kartika Maulida Hindrayani, S.Kom., M.Kom	Ayiolla Terza Damaliana, S.Si., M.Stat
NIP. 199209092022032009	NIP. 199408022022032015

Mengetahui,

Fakultas Ilmu Komputer	Program Studi Sains Data
Dekan,	Fakultas Ilmu Komputer
	Koordinator,
Prof. Dr. Ir. Novirina Hendrasarie, MT	Dr. Eng. Ir. Dwi Arman Prasetya, ST., MT., IPU
NIP. 196811261994032001	NIP. 198012052005011002

Surabaya, Juli, 2024

SURAT PERNYATAAN

Saya yang bertanda tangan di bawah ini:

Nama	: Brescia Ayundina Yuniarossy
NPM	: 20083010021
Program Studi	: Sains Data

Menyatakan bahwa judul Skripsi / Tugas Akhir sebagai berikut:

Implementasi Model *IndoBERT* Pada Analisis Sentimen Terhadap Isu Feminisme di Twitter

Bukan merupakan plagiat dari Skripsi/ Tugas Akhir/ Penelitian orang lain dan juga bukan merupakan produk/ *software*/ hasil karya yang saya beli dari orang lain

Saya juga menyatakan bahwa Skripsi/ Tugas Akhir ini adalah pekerjaan saya sendiri, kecuali yang dinyatakan dalam Daftar Pustaka, dan tidak pernah diajukan untuk syarat memperoleh gelar di Universitas Pembangunan Nasional "Veteran" Jawa Timur maupun di institusi pendidikan lain.
Jika ternyata dikemudian hari pernyataan ini terbukti tidak benar, maka Saya bertanggung jawab penuh dan siap menerima segala konsekuensi, termasuk pembatalan ijazah dikemudian hari

Surabaya, 05 Juli 2024

Hormat Saya

Brescia Ayundina Yuniarossy
NPM. 20083010021

ABSTRAK

IMPLEMENTASI MODEL *INDOBERT* PADA ANALISIS SENTIMEN TERHADAP ISU FEMINISME DI TWITTER

Nama Mahasiswa / NPM	**:**	**Brescia Ayundina Yuniarossy / 20083010021**
Program Studi	**:**	**Sains Data, FASILKOM,UPN Veteran Jatim**
Dosen Pembimbing 1	**:**	**Kartika Maulida Hindrayani, S.Kom., M.Kom**
Dosen Pembimbing 2	**:**	**Aviolla Terza Damaliana, S.Si., M.Stat**

ABSTRAK

Penelitian ini berfokus pada analisis sentimen masyarakat terhadap isu Kekerasan Dalam Rumah Tangga (KDRT) dan Pelecehan Seksual di Indonesia, dua isu sosial yang serius dan berdampak luas. Dalam era digital, media sosial seperti Twitter menjadi *platform* untuk mengekspresikan opini publik. Menggunakan model *IndoBERT*, varian dari *Bidirectional Encoder Representations from Transformers* (BERT) khusus bahasa Indonesia untuk analisis sentimen masyarakat terhadap dua isu sosial. Penelitian ini menunjukkan bahwa *IndoBERT* mampu mengklasifikasikan sentimen publik dengan akurasi 89% pada dataset Pelecehan Seksual tanpa ekstrasi fitur *Bag of Words* (BoW) dan SMOTE. Data yang digunakan dalam penelitian ini didapatkan melalui proses *crawling* dari Twitter dengan total 3007 data dari kedua topik tersebut. Setelah itu, dilakukan tahap *preprocessing* untuk membersihkan data dari *noise*. Sentimen dari *tweet* yang terkumpul diberikan label positif, negatif, dan netral. Penelitian ini menunjukkan bahwa *IndoBERT* efektif dalam mengklasifikasikan sentimen publik dan dapat membantu dalam memahami persepsi masyarakat terhadap isu-isu feminisme, khususnya KDRT dan pelecehan seksual.

Kata Kunci: Analisis Sentimen, Feminisme, *IndoBERT*, *BoW*, *SMOTE*

ABSTRACT

IMPLEMENTATION OF *INDOBERT* MODEL ON SENTIMENT ANALYSIS OF FEMINISM ISSUES ON TWITTER

Student Name / NPM	**:**	**Brescia Ayundina Yuniarossy / 20083010021**
Study Program	**:**	**Sains Data, FASILKOM,UPN Veteran Jatim**
Advisor 1	**:**	**Kartika Maulida Hindrayani, S.Kom., M.Kom**
Advisor 2	**:**	**Aviolla Terza Damaliana, S.Si., M.Stat**

ABSTRACT

This research focuses on analyzing public sentiment towards the issues of Domestic Violence and Sexual Harassment in Indonesia, two serious and far-reaching social issues. In the digital age, social media such as Twitter has become a platform for expressing public opinion. Using the IndoBERT model, a variant of the Indonesian language-specific Bidirectional Encoder Representations from Transformers (BERT) for public sentiment analysis on two social issues. This research shows that IndoBERT is able to classify public sentiment with 89% accuracy on the Sexual Harassment dataset without Bag of Words (BoW) and SMOTE feature extraction. The data used in this study was obtained through a crawling process from Twitter with a total of 3007 data from both topics. After that, a preprocessing stage is carried out to clean the data from noise. The sentiments of the collected tweets are labeled as positive, negative, and neutral. This research shows that IndoBERT is effective in classifying public sentiment and can help in understanding the public perception of feminism issues, especially domestic violence and sexual harassment.

Keywords: Sentiment Analysis, Feminism, IndoBERT, BoW, SMOTE

KATA PENGANTAR

Puji dan syukur kehadirat ALLAH SWT, atas limpahan Rahmat serta Kasih Sayang-Nya sehingga penulis dapat menyelesaikan Proposal Skripsi yang merupakan persyaratan dalam menyelesaikan mata kuliah Seminar Proposal pada Program Studi S1 Sains Data di Universitas Pembangunan Nasional "Veteran" Jawa Timur.

Dalam penyusunan Skripsi ini tidak terlepas dari bantuan berbagai pihak, dan dalam kesempatan ini penulis ingin mengucapkan terima kasih kepada:

1. Kedua orang tua dan keluarga yang selalu memberikan doa dan dukungan kepada penulis, sehingga penulis bersemangat dalam menyelesaikan skripsi ini.
2. Ibu Kartika Maulida Hindrayani, S.Kom., M.Kom., selaku dosen pembimbing 1 penulis selama penelitian ini dilakukan.
3. Ibu Aviolla Terza Damaliana, S.Si., M.Stat., selaku dosen pembimbing 2 penulis selama penelitian ini dilakukan.
4. Bapak Dr. Ir. Mohammad Idhom, SP., S.Kom., MT, selaku dosen penguji 1 penulis yang sangat baik dan mengarahkan penulis selama proses revisi berlangsung.
5. Bapak Trimono, S.Si., M.Si, selaku dosen penguji 2 penulis yang mampu mengarahkan penulis dalam kepenulisan skripsi ini.
6. Bapak Tresna Maulana Fahrudin, S.ST., M.T., selaku dosen wali penulis.
7. Kepada diri sendiri (penulis) terima kasih karena kamu sudah bertahan dan mampu menyelesaikan skripsi ini dengan sehat dan semangat juang.
8. Brescia Ayundini Yuniarossy selaku kembaran perempuan yang telah memberikan dukungan dan tahap proses skripsi ini.
9. Rangga Saputra selaku teman penulis yang memiliki peranan penting selama penelitian ini.
10. Alm. Muhammad Jibril Izzah Nur Rahm selaku teman penulis yang sangat berharga.

11. Kepada All Unit NCT, Red Velvet, Aespa, iKON, BlackPink, BABYMONSTER yang sudah mendukung penulis (secara tidak langsung) dengan karyanya yang menjadi penyemangat dalam mengerjakan skripsi ini.
12. Lee Taeyong, Lee Jeno, Nakamoto Yuta, Winwin, Xiaojun, Koo Junhoe, Kang Seulgi, NingNing, Kwon Yuri, Kyuhyun, Kang Daniel, Ong Seongwu, Park Woojin, Kim Jaehwan yang telah memberikan gambaran perjuangan sampai titik sukses kepada penulis agar cepat menyelesaikan skripsi ini.
13. Kepada Nurhalimah selaku teman *online* yang saya sayangi dan selalu berada disamping saya sedari saya SMP.
14. Kepada Dea Salfia Murdiasari selaku teman kost saya yang selalu berada disamping bahkan menemani penulis selama proses pembuatan skripsi ini.
15. Teman seperjuangan penulis selama menempuh Pendidikan di Sains Data.

Penulis menyadari bahwa masih terdapat banyak kekurangan dalam Proposal Skripsi ini, namun penulis berharap semoga Proposal Skripsi ini dapat memberikan kontribusi terhadap perkembangan ilmu pengetahuan, khususnya dalam bidang ilmu Sains Data.

Surabaya, 15 Juli 2024

Penulis

DAFTAR ISI

DAFTAR GAMBAR

DAFTAR TABEL

LAMPIRAN