

**DETEKSI *TWEET* PELECEHAN SEKSUAL DI MEDIA SOSIAL
TWITTER MENGGUNAKAN *MACHINE LEARNING***

TUGAS AKHIR

Oleh:

MUHAMMAD AFRIZAL MIQDAD

NIM. 1910817210011



**PROGRAM STUDI TEKNOLOGI INFORMASI
FAKULTAS TEKNIK
UNIVERSITAS LAMBUNG MANGKURAT
BANJARMASIN
2023**

**DETEKSI *TWEET* PELECEHAN SEKSUAL DI MEDIA SOSIAL
TWITTER MENGGUNAKAN *MACHINE LEARNING***

TUGAS AKHIR

Diajukan Untuk Memenuhi Salah Satu Syarat
Sarjana Strata-1 Teknologi Informasi

Oleh:

MUHAMMAD AFRIZAL MIQDAD

NIM. 1910817210011



**PROGRAM STUDI TEKNOLOGI INFORMASI
FAKULTAS TEKNIK
UNIVERSITAS LAMBUNG MANGKURAT
BANJARMASIN, JUNI 2023**

LEMBAR PERNYATAAN

Saya yang bertanda tangan di bawah ini,

Nama : Muhammad Afrizal Miqdad
NIM : 1910817210011
Fakultas : Teknik
Program Studi : Teknologi Informasi
Judul Tugas Akhir : Deteksi *Tweet* Pelecehan Seksual di Media Sosial
Twitter Menggunakan *Machine Learning*
Pembimbing : Dr. Ir. Yuslena Sari, S.Kom, M.Kom.

Dengan ini menyatakan bahwa dalam Tugas Akhir ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar akademik di suatu perguruan tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar rujukan.

Banjarmasin, 12 Juni 2023

Muhammad Afrizal Miqdad

1910817210011

LEMBAR PENGESAHAN
SKRIPSI PROGRAM STUDI S-1 TEKNOLOGI INFORMASI
Deteksi Tweet Pelecehan Seksual di Media Sosial Twitter Menggunakan
Machine Learning
oleh
Muhammad Afrizal Miqdad (1910817210011)

Telah dipertahankan di depan Tim Penguji pada 21 Juni 2023 dan
dinyatakan

L U L U S

Komite Penguji :

Ketua : Eka Setya Wijaya, S.T., M.Kom.
NIP 198205082008011010

Anggota 1 : Nurul Fathanah Mustamin, S.Pd., M.T.
NIP 199110252019032018

Anggota 2 : Andreyan Rizky Baskara, S.Kom., M.Kom.
NIP 199307032019031011

Pembimbing Utama : Dr. Ir. Yuslena Sari, S.Kom., M.Kom.
NIP 198411202015042002



Banjarbaru, 22 JUN 23
diketahui dan disahkan oleh:

Wakil Dekan Bidang Akademik
Fakultas Teknik ULM,



Muhammad, S.T., M.T.
NIP 197401071998021001

Koordinator Program Studi
S-1 Teknologi Informasi,



Dr. Ir. Yuslena Sari, S.Kom., M.Kom.
NIP 198411202015042002

PERSETUJUAN TUGAS AKHIR

**DETEKSI *TWEET* PELECEHAN SEKSUAL DI MEDIA SOSIAL TWITTER
MENGUNAKAN *MACHINE LEARNING***

**OLEH
MUHAMMAD AFRIZAL MIQDAD
NIM. 1910817210011**

**Telah diperiksa dan terpenuhi semua persyaratan akademik, administrasi, dan
disetujui untuk dipertahankan dihadapan dewan penguji**

**Banjarmasin, 9 Juni 2023
Pembimbing Utama,**



**Dr. Ir. Yuslena Sari S.Kom., M.Kom.
NIP. 198411202015042002**

ABSTRAK

Kasus pelecehan seksual marak terjadi di media sosial akibat efek kebebasan berpendapat. Di Indonesia, kasus pelecehan seksual *online* meningkat sebesar 83% dari tahun 2020 ke tahun 2021. Twitter, sebagai salah satu platform media sosial terkemuka, sering digunakan sebagai media untuk pelecehan seksual. Sejak tahun 2012 hingga 2021, Twitter telah menanggukkan 14 juta akun pengguna dengan berbagai alasan, termasuk pelecehan seksual, dengan jumlah akun yang ditanggukkan untuk alasan ini melebihi 250 ribu. Diperlukan metode untuk menyaring teks yang dikirim di media sosial guna mengurangi jumlah kasus pelecehan seksual secara *online*. *Natural Language Processing*, yang merupakan cabang dari *machine learning*, dapat digunakan untuk mengklasifikasikan teks terkait pelecehan seksual di media sosial. Model-model *State-of-the-art* seperti BERT, RoBERTa, XLNet, IndoBERT, dan IndoXLNet telah menunjukkan performa yang baik dalam tugas klasifikasi teks. Dengan menggunakan dataset sebanyak 30.189 data yang terbagi menjadi tiga bagian: data *training*, data *validation*, dan data *testing* dengan rasio 7:3:1, serta *batch size* 16 dan *epoch* 5, model IndoBERT menghasilkan performa terbaik dengan nilai *accuracy*, *f1-score*, *recall*, dan *precision* sebesar 0,8587, 0,8646, 0,8897, dan 0,8409, secara berturut-turut. Model IndoBERT berhasil mengklasifikasikan dengan benar 1.350 data pelecehan seksual dari total 1.513 data pelecehan seksual pada data *testing*.

Kata kunci: Pelecehan Seksual, Klasifikasi, *Natural Language Processing*

ABSTRACT

Cases of sexual harassment have been rampant on social media due to the effects of freedom of expression. In Indonesia, online sexual harassment cases increased by 83% from 2020 to 2021. Twitter, as one of the leading social media platforms, is often used as a medium for sexual harassment. From 2012 to 2021, Twitter has suspended 14 million user accounts for various reasons, including sexual harassment, with the number of suspended accounts for this reason exceeding 250,000. A method is needed to filter the text sent on social media to reduce the number of online sexual harassment cases. Natural Language Processing, which is a branch of machine learning, can be used to classify text related to sexual harassment on social media. State-of-the-art models such as BERT, RoBERTa, XLNet, IndoBERT, and IndoXLNet have shown good performance in text classification tasks. Using a dataset of 30,189 samples divided into three parts: training, validation, and testing data with a ratio of 7:3:1, and a batch size of 16 and 5 epochs, the IndoBERT model achieved the best performance with accuracy, F1-score, recall, and precision values of 0.8587, 0.8646, 0.8897, and 0.8409, respectively. The IndoBERT model successfully classified 1,350 instances of sexual harassment correctly out of the 1,513 instances in the testing data.

Keyword: Sexual Harassment, Classification, Natural Language Processing

LEMBAR PERSEMBAHAN

Penulis mempersembahkan Tugas Akhir ini kepada:

1. Ayah, Ibu, dan seluruh keluarga tercinta yang telah memberikan semangat, motivasi, maupun bantuan berupa tenaga dan materi sehingga penulis dapat menyelesaikan perkuliahan dan penulisan Tugas Akhir ini.
2. Ibu Dr. Ir. Yuslena Sari, S.Kom., M.Kom., selaku Koordinator Program Studi Teknologi Informasi dan Dosen Pembimbing Utama yang selalu menyempatkan waktunya untuk memberikan bimbingan, arahan, saran dan dukungan bagi penulis dari awal hingga akhir yang sangat membantu dalam perkuliahan maupun penyelesaian Tugas Akhir ini.
3. Bapak Andry Fajar Zulkarnain, S.ST., M.T., selaku Dosen Pembimbing Akademik yang selalu membantu, membimbing, dan menyempatkan waktu untuk memberikan dukungan dan arahan dalam menyelesaikan perkuliahan dan Tugas Akhir ini.
4. Seluruh Dosen beserta Staf Program Studi Teknologi Informasi yang turut membantu dan mengarahkan dalam perkuliahan dan penyelesaian Tugas Akhir ini.
5. Nurul Hikmah selaku rekan yang turut andil dalam pengolahan data sehingga Tugas Akhir ini dapat selesai dengan baik.
6. Teman seperjuangan yang selalu memotivasi, mendukung, bahkan selalu ikhlas dan sigap membantu dalam perkuliahan hingga Tugas Akhir, Digdo Aji Asrowi, Irvan Aulia Luthfi, Muhammad Ardhy Satrio Jati, Muhammad Arras, Muhammad Fachrurrazi, Muhammad Faidhorrahman, Muhammad Miftachul Falah, Muhammad Nur Abdi, dan Nurul Hikmah serta seluruh teman-teman Angkatan 2019 Program Studi Teknologi Informasi.
7. Terakhir, diri sendiri yang telah bertahan dalam kondisi apapun hingga akhirnya tiba di titik ini.

KATA PENGANTAR

Puji syukur kehadiran Allah SWT, Tuhan Yang Maha Pengasih dan Penyayang yang telah melimpahkan Rahmat dan hidayah-Nya, yang telah memberikan kita berbagai macam nikmat dan rezeki, sehingga semua cita-cita serta harapan yang ingin kita capai menjadi lebih mudah. Sholawat serta salam tidak lupa penulis sampaikan kepada junjungan Nabi Besar Muhammad SAW yang telah membawa kita ke jalan yang terang benderang sehingga penulis dapat menyelesaikan Tugas Akhir ini dengan Judul “Deteksi *Tweet* Pelecehan Seksual di Media Sosial Twitter Menggunakan *Machine Learning*”. Penulis menyadari bahwa penyelesaian laporan ini tidak akan terwujud tanpa adanya bimbingan dan bantuan dari berbagai pihak, dalam kesempatan ini penulis menyampaikan terimakasih yang sebesar-besarnya kepada:

1. Rektor Universitas Lambung Mangkurat, Bapak Prof. Dr. Ahmad, S.E., M.Si., yang memimpin dan memajemen jalannya seluruh perkuliahan yang ada di Universitas Lambung Mangkurat.
2. Dekan Fakultas Teknik, Bapak Prof. Dr. Ir. Irphan Fitriani Radam, S.T., M.T., IPU, yang memberikan layanan terbaik dalam perkuliahan, terkhusus pada pelaksanaan Tugas Akhir di lingkungan Fakultas Teknik.
3. Koordinator Program Studi Teknologi Informasi serta Pembimbing Utama, Ibu Dr. Ir. Yuslena Sari, S.Kom., M.Kom. yang telah memberikan arahan, bimbingan, serta solusi dalam penyelesaian Tugas Akhir.
4. Bapak Ahsani Taqwiem, S.Pd., M.Pd., yang telah memberikan masukan dan bantuan kepada penulis dalam menyelesaikan tugas akhir dari aspek kebahasaan
5. Dosen-dosen beserta staff di Program Studi Teknologi Informasi yang telah mengarahkan dan teman-teman yang membantu dalam proses penyelesaian Tugas Akhir.

Akhir kata, penulis menyampaikan terima kasih kepada semua pihak yang turut serta membantu dalam penyelesaian laporan Tugas Akhir ini. Harapan terbesar dari penyusunan laporan ini adalah dapat bermanfaat baik untuk pembaca serta mendapatkan saran dan kritik demi perbaikan dan penyempurnaan laporan ini.

Banjarmasin, 12 Juni 2023

Penulis,

Muhammad Afrizal Miqdad

DAFTAR ISI

HALAMAN SAMPUL LUAR	i
HALAMAN SAMPUL DALAM	ii
LEMBAR PERNYATAAN	iii
LEMBAR PENGESAHAN	iv
LEMBAR PERSETUJUAN.....	v
ABSTRAK	vi
ABSTRACT.....	vii
LEMBAR PERSEMBAHAN	viii
KATA PENGANTAR	ix
DAFTAR ISI.....	xi
DAFTAR TABEL.....	xiv
DAFTAR GAMBAR	xvii
DAFTAR LAMPIRAN	xx
DAFTAR RIWAYAT HIDUP.....	xxi
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Batasan Masalah.....	4
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian.....	5
BAB II TINJAUAN PUSTAKA.....	6
2.1 Landasan Teori	6
2.1.1 Pelecehan Seksual	6
2.1.2 <i>Natural Language Processing</i>	6
2.1.3 <i>Text Classification</i>	7
2.1.4 <i>Bidirectional Encoder Representations from Transformers (BERT)</i> 8	
2.1.5 <i>Robustly optimized BERT approach (RoBERTa)</i>	10
2.1.6 IndoBERT	10
2.1.7 XLNet.....	11
2.1.8 IndoXLNet	12
2.1.9 <i>Confusion Matrix</i>	12

2.2	Penelitian Terkait	13
2.2.1	<i>But how robust is RoBERTa actually?: A Benchmark of SOTA Transformer Networks for Sexual Harassment Detection on Twitter</i>	14
2.2.2	<i>Natural Language Processing through BERT for Identifying Gender-Based Violence Messages on Social Media</i>	14
2.2.3	<i>XLNet: Generalized Autoregressive Pretraining for Language Understanding</i>	14
2.2.4	<i>Classification of Censored Tweets in Chinese Language using XLNet</i> 15	
2.2.5	<i>Comparison of LSTM, SVM, and Naive Bayes for Classifying Sexual Harassment Tweets</i>	15
2.2.6	Deteksi Penggunaan Kalimat <i>Abesif</i> Pada Teks Bahasa Indonesia Menggunakan Metode IndoBERT	16
2.2.7	IndoXLNet: <i>Pre-trained</i> Language Model for Bahasa Indonesia... 16	
2.2.8	<i>Fine-tune BERT to Classify Hate Speech in Hindi English Code Mixed Text</i>	17
2.3	Kerangka Pemikiran	21
2.3.1	<i>Indicators</i>	21
2.3.2	<i>Proposed Method</i>	21
2.3.3	<i>Objective</i>	21
2.3.4	<i>Measurement</i>	22
BAB III METODOLOGI PENELITIAN.....		23
3.1	Alat dan Bahan Penelitian	23
3.1.1	Alat Penelitian	23
3.1.2	Bahan Penelitian.....	23
3.2	Alur Penelitian.....	24
3.2.1	Identifikasi Masalah	25
3.2.2	Studi Literatur	25
3.2.3	Pengumpulan data	25
3.2.4	Melakukan Eksperimen dan Penelitian.....	26
3.2.5	Analisa dan Pembahasan.....	30
3.2.6	Implementasi Sistem	31
3.2.7	Kesimpulan dan Saran.....	31
BAB IV HASIL DAN PEMBAHASAN		33
4.1.	Pengumpulan Data	33

4.2.	<i>Text Preprocessing</i>	34
4.2.1	<i>Case Folding dan Filtering</i>	34
4.2.2	<i>Tokenization</i>	35
4.2.3	<i>Normalization</i>	35
4.3.	Pelabelan Data	35
4.4.	Pembagian Data.....	37
4.5.	Implementasi Model.....	38
4.6.	Eksperimen Model	41
4.6.1	Parameter yang Digunakan	41
4.6.2	Eksperimen Model BERT	42
4.6.3	Eksperimen Model RoBERTa.....	50
4.6.4	Eksperimen Model IndoBERT.....	58
4.6.5	Eksperimen Model XLNet	66
4.6.6	Eksperimen Model IndoXLNet.....	74
4.7.	Pengujian Model.....	82
4.7.1	Pengujian Model BERT	82
4.7.2	Pengujian Model RoBERTa.....	83
4.7.3	Pengujian Model IndoBERT.....	84
4.7.4	Pengujian Model XLNet	85
4.7.5	Pengujian Model IndoXLNet.....	87
4.8.	Evaluasi Model Terbaik	88
4.8.1	Model BERT Terbaik.....	88
4.8.2	Model RoBERTa Terbaik	89
4.8.3	Model IndoBERT Terbaik	90
4.8.4	Model XLNet Terbaik.....	91
4.8.5	Model IndoXLNet Terbaik	92
4.8.	Analisa dan Pembahasan	93
4.9.	Implementasi Model.....	94
BAB V KESIMPULAN DAN SARAN.....		95
5.1	Kesimpulan.....	95
5.2	Saran	96
DAFTAR PUSTAKA		97
LAMPIRAN.....		102

DAFTAR TABEL

Tabel 2. 1 Tabel Confusion Matrix	13
Tabel 2. 2 Ringkasan Penelitian Terkait	18
Tabel 3. 1 Alat Penunjang Penelitian	23
Tabel 3. 2 Bahan Penelitian	23
Tabel 3. 3 Contoh Data	26
Tabel 3. 4 Proses Case Folding	28
Tabel 3. 5 Proses Punctuation Removal.....	29
Tabel 3. 6 Proses Normalization	29
Tabel 3. 7 Proses Tokenization	29
Tabel 4. 1 Kata Kunci Pengumpulan Data.....	33
Tabel 4. 2 Perbandingan Data Sebelum dan Sesudah Case Folding dan Filtering 34	
Tabel 4. 3 Perbandingan Data Sebelum dan Sesudah Tokenizing.....	35
Tabel 4. 4 Perbandingan Data Sebelum dan Sesudah Normalisasi.....	35
Tabel 4. 5 Acuan Pelabelan Data	36
Tabel 4. 6 Contoh Data Berlabel.....	36
Tabel 4. 7 Pembagian Data rasio 7:2:1	38
Tabel 4. 8 Pembagian Data Rasio 8:1:1	38
Tabel 4. 9 Pembagian Data Rasio 9:0.5:0.5	38
Tabel 4. 10 Model dan Sumber	41
Tabel 4. 11 Parameter yang tidak Diubah.....	41
Tabel 4. 12 Parameter yang Diubah.....	42
Tabel 4. 13 Hasil Eksperimen Pertama Model BERT	42
Tabel 4. 14 Hasil Eksperimen Kedua Model BERT	43
Tabel 4. 15 Hasil Eksperimen Ketiga Model BERT.....	45
Tabel 4. 16 Hasil Eksperimen Keempat Model BERT	46
Tabel 4. 17 Hasil Eksperimen Kelima Model BERT.....	47
Tabel 4. 19 Hasil Eksperimen Keenam Model BERT	49
Tabel 4. 20 Hasil Eksperimen Pertama Model RoBERTa.....	50
Tabel 4. 21 Hasil Eksperimen Kedua Model RoBERTa.....	51
Tabel 4. 22 Hasil Eksperimen Ketiga Model RoBERTa	53

Tabel 4. 23 Hasil Eksperimen Keempat Model RoBERTa.....	54
Tabel 4. 24 Hasil Eksperimen Kelima Model RoBERTa	55
Tabel 4. 26 Hasil Eksperimen Keenam Model RoBERTa.....	57
Tabel 4. 27 Hasil Eksperimen Pertama Model IndoBERT	58
Tabel 4. 28 Hasil Eksperimen Kedua Model IndoBERT.....	59
Tabel 4. 29 Hasil Eksperimen Ketiga Model IndoBERT	61
Tabel 4. 30 Hasil Eksperimen Keempat Model IndoBERT.....	62
Tabel 4. 31 Hasil Eksperimen Kelima Model IndoBERT	63
Tabel 4. 33 Hasil Eksperimen Keenam Model IndoBERT	65
Tabel 4. 34 Hasil Eksperimen Pertama Model XLNet	66
Tabel 4. 35 Hasil Eksperimen Kedua Model XLNet	67
Tabel 4. 36 Hasil Eksperimen Ketiga Model XLNet.....	69
Tabel 4. 37 Hasil Eksperimen Keempat Model XLNet	70
Tabel 4. 38 Hasil Eksperimen Kelima Model XLNet.....	71
Tabel 4. 40 Hasil Eksperimen Keenam Model XLNet	73
Tabel 4. 41 Hasil Eksperimen Pertama Model IndoXLNet	74
Tabel 4. 42 Hasil Eksperimen Kedua Model IndoXLNet.....	75
Tabel 4. 43 Hasil Eksperimen Ketiga Model IndoXLNet.....	77
Tabel 4. 44 Hasil Eksperimen Keempat Model IndoXLNet.....	78
Tabel 4. 45 Hasil Eksperimen Kelima Model IndoXLNet	79
Tabel 4. 47 Hasil Eksperimen Keenam Model IndoXLNet.....	81
Tabel 4. 48 Hasil Pengujian Model BERT.....	82
Tabel 4. 49 Hasil Pengujian Model RoBERTa	83
Tabel 4. 50 Hasil Pengujian Model IndoBERT	84
Tabel 4. 51 Hasil Pengujian Model XLNet.....	86
Tabel 4. 52 Hasil Pengujian Model IndoXLNet	87
Tabel 4. 53 Hasil klasifikasi Data Testing Model BERT	88
Tabel 4. 54 Data yang Diklasifikasikan dengan Benar	88
Tabel 4. 55 Data yang Salah Diklasifikasikan	88
Tabel 4. 56 Hasil klasifikasi Data Testing Model RoBERTa	89
Tabel 4. 57 Data yang Diklasifikasikan dengan Benar	89
Tabel 4. 58 Data yang Salah Diklasifikasikan	89

Tabel 4. 59 Hasil klasifikasi Data Testing Model IndoBERT	90
Tabel 4. 60 Data yang Diklasifikasikan dengan Benar	90
Tabel 4. 61 Data yang Salah Diklasifikasikan	90
Tabel 4. 62 Hasil klasifikasi Data Testing Model XLNet.....	91
Tabel 4. 63 Data yang Diklasifikasikan dengan Benar	91
Tabel 4. 64 Data yang Salah Diklasifikasikan	91
Tabel 4. 65 Hasil klasifikasi Data Testing Model IndoXLNet	92
Tabel 4. 66 Data yang Diklasifikasikan dengan Benar	92
Tabel 4. 67 Data yang Salah Diklasifikasikan	92
Tabel 4. 68 Perbandingan Performa Setiap Model Terbaik.....	93

DAFTAR GAMBAR

Gambar 1. 1 Alasan-alasan Penangguhan Akun Twitter Pada Tahun 2021 [7].....	2
Gambar 2. 1 Empat Fase Text Classification [20]	8
Gambar 2. 2 Gambaran Proses <i>Pre-Training</i> dan Fine-Tuning BERT [21]	9
Gambar 2. 3 Gambaran dari arsitektur Two-Stream Self-Attention dan Permutation Language Modeling [13]	12
Gambar 2. 4 Kerangka Pemikiran Penelitian	21
Gambar 3. 1 Alur Penelitian.....	24
Gambar 3. 2 Tahapan Eksperimen dan Penelitian	27
Gambar 3. 3 Skema Eksperimen dan Pengujian	30
Gambar 4. 1 Ilustrasi Proses Scrapping	34
Gambar 4. 2 Frekuensi Kata yang Paling Muncul Perlabel	37
Gambar 4. 3 Grafik Eksperimen Pertama Model BERT Batch size 16.....	42
Gambar 4. 4 Grafik Eksperimen Pertama Model BERT Batch size 32.....	43
Gambar 4. 5 Grafik Eksperimen Kedua Model BERT Batch size 16.....	44
Gambar 4. 6 Grafik Eksperimen Kedua Model BERT Batch size 32.....	44
Gambar 4. 7 Grafik Eksperimen Ketiga Model BERT Batch Size 16.....	45
Gambar 4. 8 Grafik Eksperimen Ketiga Model BERT Batch Size 32.....	45
Gambar 4. 9 Grafik Eksperimen Keempat Model BERT Batch size 16.....	46
Gambar 4. 10 Grafik Eksperimen Keempat Model BERT Batch size 32.....	47
Gambar 4. 11 Grafik Eksperimen Kelima Model BERT Batch Size 16.....	48
Gambar 4. 12 Grafik Eksperimen Kelima Model BERT Batch Size 32.....	48
Gambar 4. 13 Grafik Eksperimen Keenam Model BERT Batch size 16.....	49
Gambar 4. 14 Grafik Eksperimen Keenam Model BERT Batch size 32.....	49
Gambar 4. 15 Grafik Eksperimen Pertama Model RoBERTa Batch size 16.....	50
Gambar 4. 16 Grafik Eksperimen Pertama Model RoBERTa Batch size 32.....	51
Gambar 4. 17 Grafik Eksperimen Kedua Model RoBERTa Batch size 16	52
Gambar 4. 18 Grafik Eksperimen Kedua Model RoBERTa Batch size 32	52
Gambar 4. 19 Grafik Eksperimen Ketiga Model RoBERTa Batch Size 16	53
Gambar 4. 20 Grafik Eksperimen Ketiga Model RoBERTa Batch Size 32	53
Gambar 4. 21 Grafik Eksperimen Keempat Model RoBERTa Batch size 16	54

Gambar 4. 22 Grafik Eksperimen Keempat Model RoBERTa Batch size 32	55
Gambar 4. 23 Grafik Eksperimen Kelima Model RoBERTa Batch Size 16	56
Gambar 4. 24 Grafik Eksperimen Kelima Model RoBERTa Batch Size 32	56
Gambar 4. 25 Grafik Eksperimen Keenam Model RoBERTa Batch size 16	57
Gambar 4. 26 Grafik Eksperimen Keenam Model RoBERTa Batch size 32	57
Gambar 4. 27 Grafik Eksperimen Pertama Model IndoBERT Batch size 16.....	58
Gambar 4. 28 Grafik Eksperimen Pertama Model IndoBERT Batch size 32.....	59
Gambar 4. 29 Grafik Eksperimen Kedua Model IndoBERT Batch size 16	60
Gambar 4. 30 Grafik Eksperimen Kedua Model IndoBERT Batch size 32	60
Gambar 4. 31 Grafik Eksperimen Ketiga Model IndoBERT Batch Size 16	61
Gambar 4. 32 Grafik Eksperimen Ketiga Model IndoBERT Batch Size 32	61
Gambar 4. 33 Grafik Eksperimen Keempat Model IndoBERT Batch size 16	62
Gambar 4. 34 Grafik Eksperimen Keempat Model IndoBERT Batch size 32	63
Gambar 4. 35 Grafik Eksperimen Kelima Model IndoBERT Batch Size 16	64
Gambar 4. 36 Grafik Eksperimen Kelima Model IndoBERT Batch Size 32	64
Gambar 4. 37 Grafik Eksperimen Keenam Model IndoBERT Batch size 16	65
Gambar 4. 38 Grafik Eksperimen Keenam Model IndoBERT Batch size 32	65
Gambar 4. 39 Grafik Eksperimen Pertama Model XLNet Batch size 16	66
Gambar 4. 40 Grafik Eksperimen Pertama Model XLNet Batch size 32	67
Gambar 4. 41 Grafik Eksperimen Kedua Model XLNet Batch size 16.....	68
Gambar 4. 42 Grafik Eksperimen Kedua Model XLNet Batch size 32.....	68
Gambar 4. 43 Grafik Eksperimen Ketiga Model XLNet Batch Size 16.....	69
Gambar 4. 44 Grafik Eksperimen Ketiga Model XLNet Batch Size 32.....	69
Gambar 4. 45 Grafik Eksperimen Keempat Model XLNet Batch size 16.....	70
Gambar 4. 46 Grafik Eksperimen Keempat Model XLNet Batch size 32.....	71
Gambar 4. 47 Grafik Eksperimen Kelima Model XLNet Batch Size 16.....	72
Gambar 4. 48 Grafik Eksperimen Kelima Model XLNet Batch Size 32.....	72
Gambar 4. 49 Grafik Eksperimen Keenam Model XLNet Batch size 16.....	73
Gambar 4. 50 Grafik Eksperimen Keenam Model XLNet Batch size 32.....	73
Gambar 4. 51 Grafik Eksperimen Pertama Model IndoXLNet Batch size 16.....	74
Gambar 4. 52 Grafik Eksperimen Pertama Model IndoXLNet Batch size 32.....	75
Gambar 4. 53 Grafik Eksperimen Kedua Model IndoXLNet Batch size 16	76

Gambar 4. 54 Grafik Eksperimen Kedua Model IndoXLNet Batch size 32	76
Gambar 4. 55 Grafik Eksperimen Ketiga Model IndoXLNet Batch Size 16.....	77
Gambar 4. 56 Grafik Eksperimen Ketiga Model IndoXLNet Batch Size 32.....	77
Gambar 4. 57 Grafik Eksperimen Keempat Model IndoXLNet Batch size 16 ...	78
Gambar 4. 58 Grafik Eksperimen Keempat Model IndoXLNet Batch size 32	79
Gambar 4. 59 Grafik Eksperimen Kelima Model IndoXLNet Batch Size 16	80
Gambar 4. 60 Grafik Eksperimen Kelima Model IndoXLNet Batch Size 32	80
Gambar 4. 61 Grafik Eksperimen Keenam Model IndoXLNet Batch size 16.....	81
Gambar 4. 62 Grafik Eksperimen Keenam Model IndoXLNet Batch size 32.....	81
Gambar 4. 63 Tampilan Antarmuka Sistem.....	95
Gambar 4. 64 Contoh Masukan Teks Pelecehan Seksual	95
Gambar 4. 65 Contoh Masukan Teks Bukan Pelecehan Seksual.....	96

DAFTAR LAMPIRAN

Lampiran 1. Lembar Konsultasi.....	102
Lampiran 2. Baris Kode Scrapping	103
Lampiran 3. Baris Kode Case Folding dan Filtering	106
Lampiran 4. Baris Kode Tokenization	107
Lampiran 5. Baris Kode Normalization	108
Lampiran 6. Kode untuk Pembagian data	111
Lampiran 7. Kode untuk Implementasi Model	112
Lampiran 8. Kuesioner Keyword Scrapping.....	118
Lampiran 9. Validasi Dataset.....	119

DAFTAR RIWAYAT HIDUP



Nama : Muhammad Afrizal Miqdad
Tempat dan Tanggal Lahir : Barabai, 26 September 2001
Alamat : Jl. Rutas, RT.09/ RW.03. Desa Kayu Bawang, Kecamatan Barabai, Kabupaten Hulu Sungai Tengah, Kalimantan Selatan, 71315
Email : afrizal.bbm123@gmail.com
No. HP : 085654051875

Agama : Islam
Kewarganegaraan : Indonesia
Nama Orang Tua : Muhammad Rizani, S.Pd.
Norhayati, S.Pd.
Anak ke- : 1 dari 2 bersaudara
Riwayat Pendidikan : SDN 1 Kayu Bawang
SMPN 1 Barabai
SMAN 1 Barabai
S-1 Teknologi Informasi Universitas Lambung Mangkurat