

**EVALUASI KOMPARATIF ALGORITMA *MACHINE LEARNING*
SVM + TF-IDF DENGAN *RANDOM FOREST + WORD*
EMBEDDING FASTTEXT UNTUK DETEKSI HOAKS DI
INDONESIA**

SKRIPSI

Oleh:

**VANESA PUTRI WIJAYA
NIM. 221130405
WELISMAN KRISTIAN MADUWU
NIM. 221130743
JERRICK EASON CHOO
NIM. 221130339**



**PROGRAM STUDI S-1 TEKNOLOGI INFORMASI
FAKULTAS INFORMATIKA
UNIVERSITAS MIKROSKIL
MEDAN
2026**

**COMPARATIVE EVALUATION OF SVM + TF-IDF MACHINE
LEARNING ALGORITHMS WITH RANDOM FOREST + WORD
EMBEDDING FASTTEXT FOR HOAX DETECTION IN
INDONESIA**

UNDERGRADUATE THESIS

By:

**VANESA PUTRI WIJAYA
ID NUMBER. 221130405
WELISMAN KRISTIAN MADUWU
ID NUMBER. 221130743
JERRICK EASON CHOO
ID NUMBER. 221130339**



**UNDERGRADUATE PROGRAM OF INFORMATION TECHNOLOGY
FACULTY OF INFORMATICS
UNIVERSITAS MIKROSKIL
MEDAN
2026**

LEMBARAN PENGESAHAN

**EVALUASI KOMPARATIF ALGORITMA *MACHINE LEARNING*
SVM+TF-IDF DENGAN *RANDOM FOREST+WORD EMBEDDING*
FASTTEXT UNTUK DETEKSI HOAKS DI INDONESIA**

SKRIPSI

Diajukan untuk Melengkapi Persyaratan Guna
Mendapatkan Gelar Sarjana Teknologi Informasi
Program Studi S-1 Teknologi Informasi

Oleh:

VANESA PUTRI WIJAYA

NIM. 221130405

WELISMAN KRISTIAN MADUWU

NIM. 221130743

JERRICK EASON CHOO

NIM. 221130339

Disetujui Oleh:

Dosen Pembimbing I,

Dosen Pembimbing II,


Carles Juliandy, S.Kom., M.Kom.


Tanti, S.E., M.Kom.

Medan, 15 Juli 2026

Diketahui dan Disahkan Oleh:

Ketua Program Studi
S-1 Teknologi Informasi,


Mustika Ulina, S.Kom. M.Kom.

HALAMAN PERNYATAAN

Saya yang membuat pernyataan ini adalah mahasiswa Program Studi S-1 Teknologi Informasi Universitas Mikroskil Medan dengan identitas mahasiswa sebagai berikut:

Nama : Vanesa Putri Wijaya

NIM : 221130405

Saya telah melaksanakan penelitian dan penulisan Skripsi Teknologi Informasi dengan judul dan tempat penelitian sebagai berikut:

Judul Skripsi : Evaluasi Komparatif Algoritma Machine Learning SVM+TFIDF dengan Random Forest+Word Embedding untuk Deteksi Hoaks di Indonesia

Tempat Penelitian : Universitas Mikroskil

Alamat Tempat Penelitian : JL. M.H Thamrin No. 140, Pusat Ps., Kec. Medan Kota, Kota Medan, Sumatera Utara.

No. Telp. Tempat Penelitian : (061) 4573767

Sehubungan dengan Skripsi Teknologi Informasi tersebut, dengan ini saya menyatakan dengan sebenar-benarnya bahwa penelitian dan penulisan Skripsi Teknologi Informasi tersebut merupakan hasil karya saya sendiri (tidak menyuruh orang lain yang mengerjakannya) dan semua sumber, baik yang dikutip maupun dirujuk, telah saya nyatakan dengan benar. Bila di kemudian hari ternyata terbukti bahwa bukan saya yang mengerjakannya (membuatnya), maka saya bersedia dikenakan sanksi yang telah ditetapkan oleh Universitas Mikroskil Medan, yakni pencabutan ijazah yang telah saya terima dan ijazah tersebut dinyatakan tidak sah.

Selain itu, demi pengembangan ilmu pengetahuan, saya menyetujui untuk memberikan kepada Universitas Mikroskil Medan Hak Bebas Royalti Non-eksklusif (Non-exclusive Royalty Free Right) atas Skripsi Teknologi Informasi saya beserta perangkat yang ada (jika diperlukan). Dengan hak ini, Universitas Mikroskil Medan berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (database), merawat, dan mempublikasikan Skripsi Teknologi Informasi saya, secara keseluruhan atau hanya sebagian atau hanya ringkasannya saja dalam bentuk format tercetak dan/atau elektronik, selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik hak cipta. Menyatakan juga bahwa saya akan mempertahankan hak eksklusif saya untuk menggunakan seluruh atau sebagian isi Skripsi Teknologi Informasi saya guna pengembangan karya di masa depan, misalnya dalam bentuk artikel, buku, ataupun perangkat lunak/sistem informasi.

Demikian pernyataan ini saya perbuat dengan sungguh-sungguh, dalam keadaan sadar dan tanpa ada tekanan dari pihak manapun.

Medan, 26 Juni 2026

Saya yang membuat pernyataan,



Vanesa Putri Wijaya

HALAMAN PERNYATAAN

Saya yang membuat pernyataan ini adalah mahasiswa Program Studi S-1 Teknologi Informasi Universitas Mikroskil Medan dengan identitas mahasiswa sebagai berikut:

Nama : Welisman Kristian Maduwu

NIM : 221130743

Saya telah melaksanakan penelitian dan penulisan Skripsi Teknologi Informasi dengan judul dan tempat penelitian sebagai berikut:

Judul Skripsi : Evaluasi Komparatif Algoritma Machine Learning SVM+TFIDF dengan Random Forest+Word Embedding untuk Deteksi Hoaks di Indonesia

Tempat Penelitian : Universitas Mikroskil

Alamat Tempat Penelitian : Jl. M.H Thamrin No. 140, Pusat Ps., Kec. Medan Kota, Kota Medan, Sumatera Utara.

No. Telp. Tempat Penelitian : (061) 4573767

Sehubungan dengan Skripsi Teknologi Informasi tersebut, dengan ini saya menyatakan dengan sebenar-benarnya bahwa penelitian dan penulisan Skripsi Teknologi Informasi tersebut merupakan hasil karya saya sendiri (tidak menyuruh orang lain yang mengerjakannya) dan semua sumber, baik yang dikutip maupun dirujuk, telah saya nyatakan dengan benar. Bila di kemudian hari ternyata terbukti bahwa bukan saya yang mengerjakannya (membuatnya), maka saya bersedia dikenakan sanksi yang telah ditetapkan oleh Universitas Mikroskil Medan, yakni pencabutan ijazah yang telah saya terima dan ijazah tersebut dinyatakan tidak sah.

Selain itu, demi pengembangan ilmu pengetahuan, saya menyetujui untuk memberikan kepada Universitas Mikroskil Medan Hak Bebas Royalti Non-eksklusif (Non-exclusive Royalty Free Right) atas Skripsi Teknologi Informasi saya beserta perangkat yang ada (jika diperlukan). Dengan hak ini, Universitas Mikroskil Medan berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (database), merawat, dan mempublikasikan Skripsi Teknologi Informasi saya, secara keseluruhan atau hanya sebagian atau hanya ringkasannya saja dalam bentuk format tercetak dan/atau elektronik, selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik hak cipta. Menyatakan juga bahwa saya akan mempertahankan hak eksklusif saya untuk menggunakan seluruh atau sebagian isi Skripsi Teknologi Informasi saya guna pengembangan karya di masa depan, misalnya dalam bentuk artikel, buku, ataupun perangkat lunak/sistem informasi.

Demikian pernyataan ini saya perbuat dengan sungguh-sungguh, dalam keadaan sadar dan tanpa ada tekanan dari pihak manapun.

Medan, 26 Juni 2026

Saya yang membuat pernyataan,



Welisman Kristian Maduwu

HALAMAN PERNYATAAN

Saya yang membuat pernyataan ini adalah mahasiswa Program Studi S-1 Teknologi Informasi Universitas Mikroskil Medan dengan identitas mahasiswa sebagai berikut:

Nama : Jerrick Eason Choo

NIM : 221130339

Saya telah melaksanakan penelitian dan penulisan Skripsi Teknologi Informasi dengan judul dan tempat penelitian sebagai berikut:

Judul Skripsi : Evaluasi Komparatif Algoritma Machine Learning SVM+TFIDF dengan Random Forest+Word Embedding untuk Deteksi Hoaks di Indonesia

Tempat Penelitian : Universitas Mikroskil

Alamat Tempat Penelitian : JL. M.H Thamrin No. 140, Pusat Ps., Kec. Medan Kota, Kota Medan, Sumatera Utara.

No. Telp. Tempat Penelitian : (061) 4573767

Sehubungan dengan Skripsi Teknologi Informasi tersebut, dengan ini saya menyatakan dengan sebenar-benarnya bahwa penelitian dan penulisan Skripsi Teknologi Informasi tersebut merupakan hasil karya saya sendiri (tidak menyuruh orang lain yang mengerjakannya) dan semua sumber, baik yang dikutip maupun dirujuk, telah saya nyatakan dengan benar. Bila di kemudian hari ternyata terbukti bahwa bukan saya yang mengerjakannya (membuatnya), maka saya bersedia dikenakan sanksi yang telah ditetapkan oleh Universitas Mikroskil Medan, yakni pencabutan ijazah yang telah saya terima dan ijazah tersebut dinyatakan tidak sah.

Selain itu, demi pengembangan ilmu pengetahuan, saya menyetujui untuk memberikan kepada Universitas Mikroskil Medan Hak Bebas Royalti Non-eksklusif (Non-exclusive Royalty Free Right) atas Skripsi Teknologi Informasi saya beserta perangkat yang ada (jika diperlukan). Dengan hak ini, Universitas Mikroskil Medan berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (database), merawat, dan mempublikasikan Skripsi Teknologi Informasi saya, secara keseluruhan atau hanya sebagian atau hanya ringkasannya saja dalam bentuk format tercetak dan/atau elektronik, selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik hak cipta. Menyatakan juga bahwa saya akan mempertahankan hak eksklusif saya untuk menggunakan seluruh atau sebagian isi Skripsi Teknologi Informasi saya guna pengembangan karya di masa depan, misalnya dalam bentuk artikel, buku, ataupun perangkat lunak/sistem informasi.

Demikian pernyataan ini saya perbuat dengan sungguh-sungguh, dalam keadaan sadar dan tanpa ada tekanan dari pihak manapun.

Medan, 26 Juni 2026

Saya yang membuat pernyataan,



Jerrick Eason Choo

EVALUASI KOMPARATIF ALGORITMA MACHINE LEARNING SVM+TF-IDF DENGAN RANDOM FOREST+WORD EMBEDDING FASTTEXT UNTUK DETEKSI HOAKS DI INDONESIA

Abstrak

Hoaks berbahasa Indonesia terus meningkat penyebarannya di media digital dan sulit dibedakan secara manual dari berita valid, sehingga diperlukan metode klasifikasi teks yang akurat untuk mendeteksinya. Penelitian ini bertujuan mengevaluasi dan membandingkan kinerja algoritma Support Vector Machine (SVM) dengan fitur Term Frequency-Inverse Document Frequency (TF-IDF) dan Random Forest dengan Word Embedding FastText dalam mendeteksi hoaks berbahasa Indonesia. Penelitian menggunakan pendekatan kuantitatif eksperimental melalui tahapan pengumpulan dataset dari portal TurnBackHoax dan Detik.com, preprocessing, ekstraksi fitur, dan pelatihan model dengan skema Stratified 5-Fold Cross Validation, serta evaluasi menggunakan metrik Accuracy, Precision, Recall, dan F1-Score. Hasil pengujian menunjukkan model Support Vector Machine (SVM) dengan Term Frequency-Inverse Document Frequency (TF-IDF) memperoleh performa terbaik dengan Accuracy 99,52%, Precision 99,54%, Recall 99,50%, dan F1-Score 99,52%, sedangkan Random Forest dengan Word Embedding FastText memperoleh Accuracy 94,18%, Precision 94,40%, Recall 93,94%, dan F1-Score 94,12%. Hasil ini meningkat signifikan dibandingkan penelitian sebelumnya, yang hanya memperoleh Accuracy 50%, Precision 51%, Recall 51%, dan F1-Score 51% untuk Support Vector Machine (SVM) dengan Term Frequency-Inverse Document Frequency (TF-IDF), serta Accuracy 70,27%, Precision 80%, Recall 54%, dan F1-Score 54% untuk Random Forest dengan Word Embedding FastText. Kombinasi SVM dan TF-IDF terbukti lebih efektif, dan penelitian ini diharapkan menjadi referensi bagi pengembangan sistem deteksi hoaks di Indonesia.

Kata kunci: Deteksi Hoaks, Support Vector Machine, Random Forest, Term Frequency-Inverse Document Frequency, Word Embedding FastText.

Abstract

Indonesian hoaxes continue to spread increasingly on digital media and are difficult to distinguish manually from valid news, requiring an accurate text classification method for detecting them. This study aims to evaluate and compare the performance of the Support Vector Machine (SVM) algorithm with Term Frequency-Inverse Document Frequency (TF-IDF) features and Random Forest with Word Embedding FastText in detecting Indonesian hoaxes. The study employs a quantitative experimental approach through stages of dataset collection from the TurnBackHoax and Detik.com portals, preprocessing, feature extraction, and model training using a Stratified 5-Fold Cross Validation scheme, with evaluation based on Accuracy, Precision, Recall, and F1-Score metrics. The experimental results show that the Support Vector Machine (SVM) with Term Frequency-Inverse Document Frequency (TF-IDF) model achieved the best performance with an Accuracy of 99,52%, Precision of 99,54%, Recall of 99,50%, and F1-Score of 99,52%, while Random Forest with Word Embedding FastText obtained an Accuracy of 94,18%, Precision of 94,40%, Recall of 93,94%, and F1-Score of 94,12%. These results show a significant improvement compared to previous research, which only achieved an Accuracy of 50%, Precision of 51%, Recall of 51%, and F1-Score of 51% for SVM with TF-IDF, as well as an Accuracy of 70.27%, Precision of 80%, Recall of 54%, and F1-Score of 54% for Random Forest with Word Embedding FastText. The combination of SVM and TF-IDF proved to be more effective, and this study is expected to serve as a reference for the development of hoax detection systems in Indonesia.

Keywords: Hoax Detection, Support Vector Machine, Random Forest, Term Frequency-Inverse Document Frequency, Word Embedding FastText.

KATA PENGANTAR

Puji dan syukur penulis panjatkan atas segala rahmat, kesehatan, serta kesempatan yang diberikan sehingga penulis dapat menyelesaikan skripsi yang berjudul “Evaluasi Komparatif Algoritma Machine Learning SVM+TF-IDF dengan Random Forest+Word Embedding FastText untuk deteksi Hoaks di Indonesia” Sebagai salah satu syarat untuk memperoleh gelar Sarjana Teknologi Informasi pada Program Studi S-1 Teknologi Informasi, Fakultas Informatika, Universitas Mikroskil.

Penyusunan skripsi ini dilakukan untuk mengevaluasi dan membandingkan performa algoritma Support Vector Machine (SVM) dengan Terms Frequency-Inverse Document Frequency (TF-IDF) serta Random Forest dengan Word Embedding FastText dalam mendeteksi berita hoaks berbahasa Indonesia. Penelitian ini diharapkan dapat memberikan kontribusi akademik sebagai referensi dalam pengembangan penelitian di bidang Machine Learning dan Natural Language Processing (NLP), khususnya pada deteksi hoaks.

Dalam proses penyusunan skripsi ini, penulis memperoleh banyak bantuan, bimbingan, dukungan, serta motivasi dari berbagai pihak. Oleh karena itu, penulis mengucapkan terima kasih kepada:

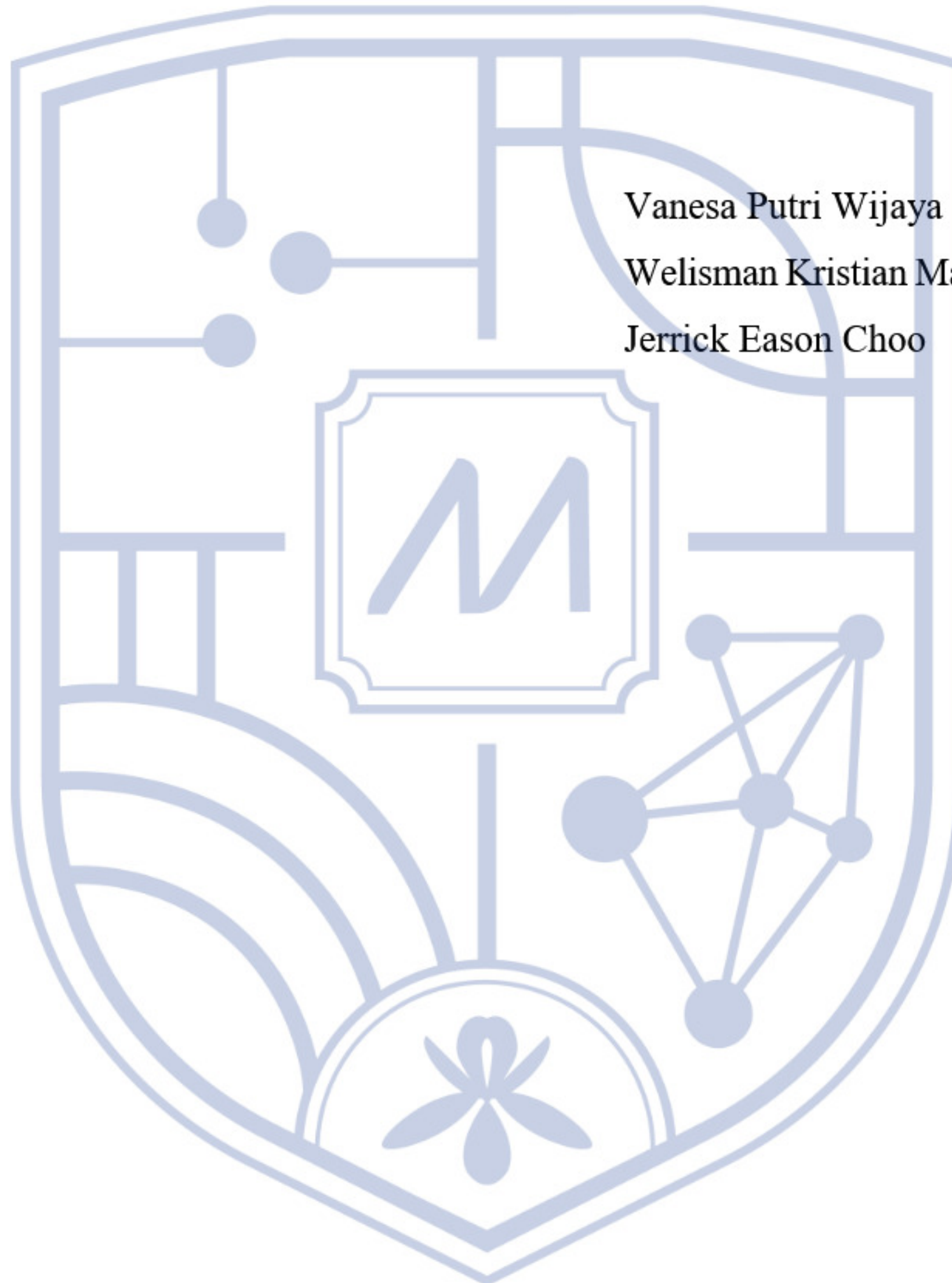
1. Bapak Carles Juliandy, S.Kom., M.Kom., selaku Dosen Pembimbing I, yang telah meluangkan waktu, memberikan arahan, bimbingan, masukan, serta motivasi yang sangat berarti sehingga penulis dapat menyelesaikan skripsi ini dengan baik.
2. Ibu Tanti, S.E., M.Kom., selaku Dosen Pembimbing II, yang telah memberikan bimbingan, saran serta masukan yang membangun selama proses penyusunan skripsi sehingga penelitian ini dapat berkembang dan diselesaikan dengan baik.
3. Bapak Hardy, S.Kom., M.Sc., Ph.D., selaku Rektor Universitas Mikroskil Medan.
4. Bapak Sunaryo Winardi, S.Kom., M.T., selaku Dekan Fakultas Informatika Universitas Mikroskil Medan.
5. Ibu Mustika Ulina, S.Kom. M.Kom., selaku Ketua Program Studi S-1 Teknologi Informasi Fakultas Informatika Universitas Mikroskil Medan.
6. Orang tua dan keluarga yang selalu memberikan doa, dukungan, semangat, dan motivasi selama proses penyusunan skripsi.
7. Seluruh dosen, staf, serta pihak-pihak lain yang telah memberikan bantuan, baik secara langsung, mau pun tidak langsung, selama proses penyusunan skripsi ini.

Penulis menyadari bahwa skripsi ini masih memiliki keterbatasan dan jauh dari sempurna. Oleh karena itu, penulis mengharapkan kritik dan saran yang membangun demi penyempurnaan penelitian di masa mendatang. Semoga skripsi ini dapat memberikan manfaat serta menjadi referensi bagi pembaca dan peneliti selanjutnya.

Medan, 26 Juni 2026

Penulis,

Vanesa Putri Wijaya
Welisman Kristian Maduwu
Jerrick Eason Choo

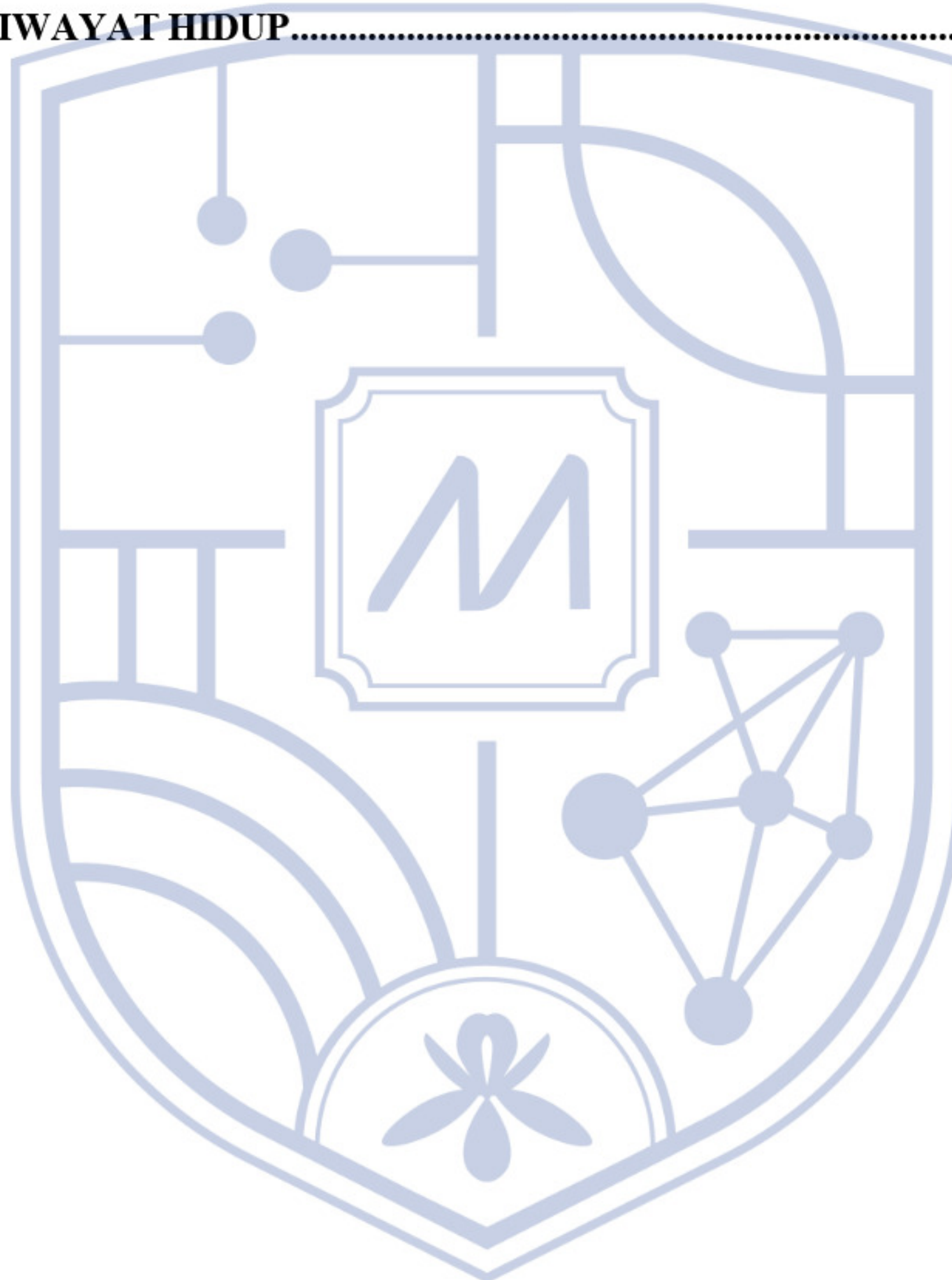


DAFTAR ISI

ABSTRAK.....	i
KATA PENGANTAR.....	ii
DAFTAR ISI.....	iv
DAFTAR GAMBAR.....	vii
DAFTAR TABEL	viii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	2
1.3 Tujuan.....	2
1.4 Manfaat.....	3
1.5 Ruang Lingkup.....	3
BAB II KAJIAN LITERATUR.....	5
2.1 Hoaks Digital	5
2.2 Natural Language Processing (NLP).....	6
2.3 Text Preprocessing.....	8
2.3.1 Case Folding.....	8
2.3.2 Tokenisasi.....	9
2.4 Ekstraksi Fitur	9
2.4.1 Term Frequency-Inverse Document Frequency (TF-IDF).....	10
2.4.2 Word Embedding FastText	11
2.5 Klasifikasi	12
2.5.1 Support Vector Machine (SVM).....	13
2.5.2 Random Forest.....	14
2.6 Pendekatan Kuantitatif Eksperimental	15
2.7 Evaluasi Model	16
2.7.1 Confusion Matrix.....	17
2.7.2 Metrik Evaluasi Klasifikasi.....	18
2.7.3 Interpretasi Nilai Evaluasi.....	22
BAB III TAHAPAN PELAKSANAAN	24
3.1 Tahapan Penelitian.....	24
3.2 Dataset Penelitian.....	25
3.3 Preprocessing	26

3.3.1 Remove Duplicate	26
3.3.2 Case Folding	28
3.3.3 Tokenisasi	29
3.3.4 Remove URL	30
3.3.5 Remove Hashtag dan Simbol.....	30
3.3.6 Stopword Removal (Bahasa Indonesia).....	32
3.3.7 Stemming (Sastrawi)	34
3.4 Pembagian Data	35
3.4.1 Stratified 5-Fold Cross Validation.....	36
3.5 Support Vector Machine (SVM)	36
3.5.1 Eksperimen Parameter TF-IDF	37
3.5.2 Ekstraksi Fitur TF-IDF	38
3.5.3 Eksperimen Parameter Support Vector Machine (SVM).....	39
3.5.4 Pelatihan Model Support Vector Machine (SVM).....	40
3.6 Random Forest.....	41
3.6.1 Ekstraksi Fitur Word Embedding FastText.....	41
3.6.2 Eksperimen Parameter Random Forest	42
3.6.3 Pelatihan Model Random Forest	44
3.7 Evaluasi Model.....	45
BAB IV HASIL DAN PEMBAHASAN	47
4.1 Hasil.....	47
4.1.1 Hasil Preprocessing.....	47
4.1.2 Hasil Pembagian Data	48
4.1.3 Hasil Implementasi Support Vector Machine Berbasis TF-IDF	48
4.1.4 Hasil Eksperimen Parameter TF-IDF.....	49
4.1.5 Hasil Ekstraksi Fitur Term Frequency-Inverse Document Frequency (TF-IDF).....	50
4.1.6 Hasil Eksperimen Parameter Support Vector Machine (SVM)	50
4.1.7 Hasil Pelatihan Model Support Vector Machine (SVM)	51
4.1.8 Hasil Evaluasi Model Support Vector Machine (SVM)	52
4.1.9 Hasil Learning Curve Support Vector Machine (SVM)	53
4.1.10 Hasil Implementasi Random Forest Berbasis Word Embedding FastText....	55
4.1.11 Hasil Ekstraksi Fitur Word Embedding FastText	55
4.1.12 Hasil Eksperimen Parameter Random Forest.....	56
4.1.13 Hasil Pelatihan Model Random Forest.....	57
4.1.14 Hasil Evaluasi Model Random Forest.....	58

4.1.15 Hasil Learning Curve Random Forest.....	59
4.2 Pembahasan.....	61
BAB V PENUTUP	64
5.1 Kesimpulan	64
5.2 Saran.....	64
DAFTAR PUSTAKA.....	66
DAFTAR RIWAYAT HIDUP.....	68
DAFTAR RIWAYAT HIDUP.....	70
DAFTAR RIWAYAT HIDUP.....	71



DAFTAR GAMBAR

Gambar 2.1 Mekanisme Klasifikasi Algoritma Random Forest [11]	14
Gambar 3.1 Tahapan Pelaksanaan	25
Gambar 3.2 Hasil Proses Remove Duplicate	27
Gambar 3.3 Pembagian Dataset	36
Gambar 3.4 Hasil Pembagian Data Setiap Fold.....	36
Gambar 3.5 Proses Eksperimen Parameter TF-IDF.....	38
Gambar 3.6 Ekstraksi Fitur TF-IDF	38
Gambar 3.7 Eksperimen Parameter SVM.....	40
Gambar 3.8 Pelatihan Model SVM.....	41
Gambar 3.9 Pemuatan Model Word Embedding FastText.....	42
Gambar 3.10 Representasi Dokumen menggunakan Mean Pooling.....	42
Gambar 3.11 Eksperimen Parameter Random Forest.....	44
Gambar 3.12 Pelatihan Model Random Forest.....	45
Gambar 4.1 Hasil Pembagian Dataset.....	48
Gambar 4.2 Hasil Eksperimen Parameter TF-IDF.....	49
Gambar 4.3 Hasil Ekstraksi Fitur TF-IDF	50
Gambar 4.4 Hasil Eksperimen Parameter SVM	51
Gambar 4.5 Hasil Prediksi Model Support Vector Machine	52
Gambar 4.6 Evaluasi Model SVM.....	52
Gambar 4.7 Classification Report SVM	52
Gambar 4.8 Analisis Kesalahan SVM	53
Gambar 4.9 Learning Curve Support Vector Machine Berbasis TF-IDF	54
Gambar 4.10 Hasil Ekstraksi Fitur Word Embedding FastText	56
Gambar 4.11 Hasil Eksperimen Parameter Random Forest	56
Gambar 4.12 Hasil Prediksi Random Forest.....	57
Gambar 4.13 Evaluasi Model Random Forest	58
Gambar 4.14 Classification Report.....	58
Gambar 4.15 Analisis Kesalahan Random Forest	59
Gambar 4.16 Learning Curve Random Forest + Word Embedding FastText	60

DAFTAR TABEL

Tabel 2.1 Tahapan Text Preprocessing dalam NLP	7
Tabel 2.2 Confusion Matrix Klasifikasi Teks Hoaks dan Non-Hoaks	18
Tabel 2.3 Interpretasi Nilai Accuracy.....	20
Tabel 2.4 Interpretasi Nilai Metrik Evaluasi Klasifikasi.....	23
Tabel 3. 1 Penghapusan Kata dan Frasa Penanda	28
Tabel 3. 2 Case Folding.....	29
Tabel 3. 3 Tokenisasi.....	29
Tabel 3. 4 Remove URL.....	30
Tabel 3. 5 Remove Hashtag dan Simbol	31
Tabel 3. 6 Stopword Removal (Bahasa Indonesia)	33
Tabel 3. 7 Stemming (Sastra)	34
Tabel 3. 8 Hyperparameter TF-IDF yang Diuji.....	37
Tabel 3. 9 Hyperparameter SVM yang Diuji	39
Tabel 3. 10 Hyperparameter Random Forest yang Diuji	43
Tabel 4. 1 Distribusi Dataset	48
Tabel 4. 2 Perbandingan Performa Model.....	61