

**ANALISIS SENTIMEN DATA PROVIDER LAYANAN
INTERNET PADA TWITTER MENGGUNAKAN *SUPPORT
VECTOR MACHINE* DENGAN PENAMBAHAN ALGORITMA
LEVENSHTEIN DISTANCE**

TESIS

Oleh

Ida Bagus Nyoman Wijana Manuaba

1729101028



**PROGRAM STUDI ILMU KOMPUTER
PROGRAM PASCASARJANA
UNIVERSITAS PENDIDIKAN GANESHA
2020**



**ANALISIS SENTIMEN DATA PROVIDER LAYANAN
INTERNET PADA TWITTER MENGGUNAKAN *SUPPORT
VECTOR MACHINE* DENGAN PENAMBAHAN ALGORITMA
LEVENSHTEIN DISTANCE**

TESIS

Diajukan kepada

Universitas Pendidikan Ganesha

Untuk Memenuhi Sebagian Persyaratan

Memperoleh Gelar Magister Komputer

Program Studi Ilmu Komputer

Oleh

Ida Bagus Nyoman Wijana Manuaba

1729101028



PROGRAM STUDI ILMU KOMPUTER

PROGRAM PASCASARJANA

UNIVERSITAS PENDIDIKAN GANESHA

2020

Tesis oleh Ida Bagus Nyoman Wijana Manuaba ini telah diperiksa dan disetujui untuk mengikuti Ujian Tesis

Singaraja, 17 Juli 2020

Pembimbing I



Dr. Gede Rasben Dantes, S.T, M.T.I

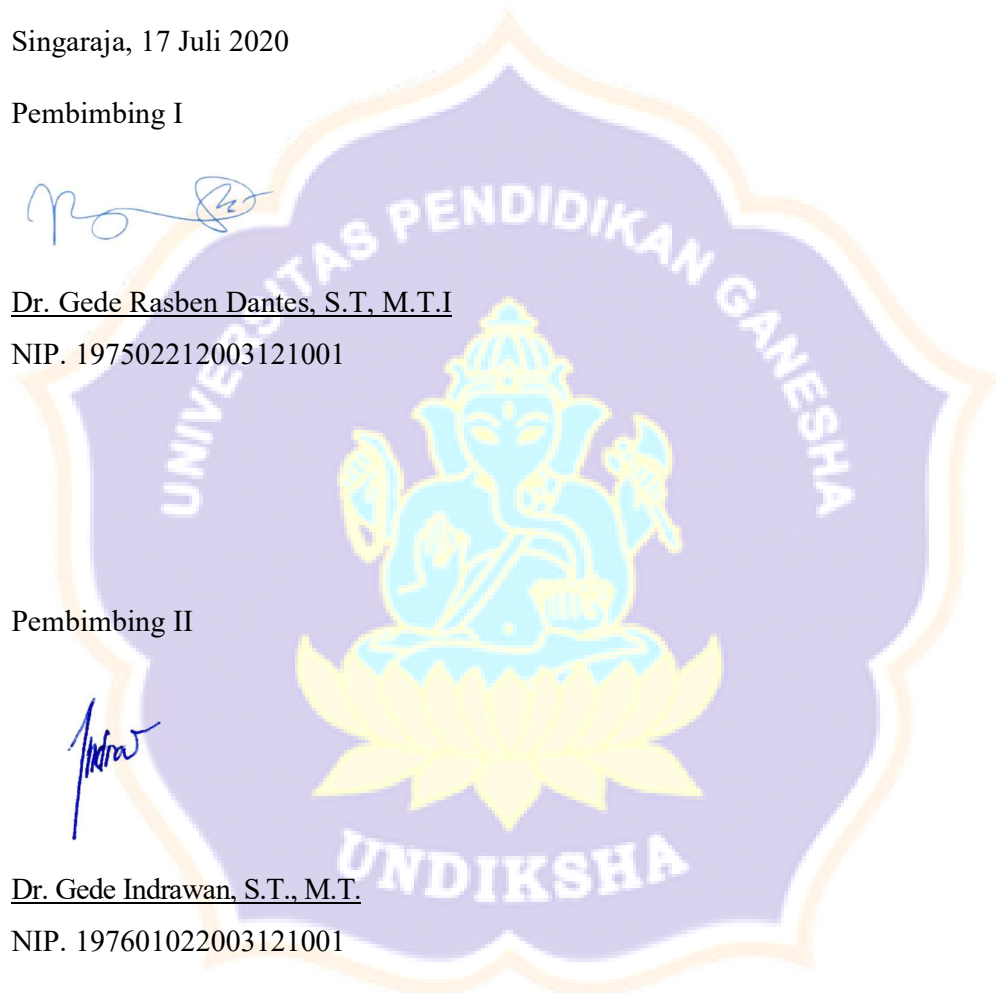
NIP. 197502212003121001

Pembimbing II



Dr. Gede Indrawan, S.T., M.T.

NIP. 197601022003121001



LEMBAR PERSETUJUAN TIM PENGUJI

Tesis oleh Ida Bagus Nyoman Wijana Manuaba ini telah berhasil dipertahankan di depan tim penguji dan dinyatakan diterima sebagai sebagian persyaratan untuk memperoleh gelar Magister Komputer di Program Studi Ilmu Komputer, Pascasarjana Universitas Pendidikan Ganesha.

Desetujui pada tanggal : 27 Juli 2020

Oleh

Tim Penguji

.....Ketua Dr. Gede Indrawan, ST.,M.T.
NIP. 19760102 200312 1 001

.....Anggota Dr. I Gede Aris Gunadi, S.Si.,M.Kom.
NIP. 19770318 200812 1 004

.....Anggota Dr. I Nyoman Sukajaya, M.T.
NIP. 19671115 199303 1 001

.....Anggota Dr. Gede Indrawan, ST.,M.T.
NIP. 19760102 200312 1 001

.....Anggota Dr. Gede Rasben Dantes, ST.,M.T.I.
NIP. 19750221 200312 1 001

Mengetahui Direktur

Pascasarjana Universitas Pendidikan Ganesha



Prof. Dr. I Gusti Putu Suharta, M.Si.

NIP. 19621215 198803 1 002

LEMBAR PERNYATAAN

Saya menyatakan dengan sesungguhnya bahwa tesis yang saya susun sebagai syarat untuk memperoleh gelar Magister Komputer dari Program Pascasarjana Universitas Pendidikan Ganesha seluruhnya merupakan hasil karya saya sendiri. Bagian-bagian tertentu dalam penulisan tesis yang saya kutip dari karya orang lain telah ditulis sumbernya secara jelas dan sesuai dengan norma, kaidah, serta etika akademis.

Apabila di kemudian hari ditemukan seluruh atau sebagian tesis ini bukan hasil karya saya sendiri atau adanya plagiat dalam bagian-bagian tertentu, saya bersedia menerima sanksi pencabutan gelar akademik yang saya sandang dan sanksi-sanksi lainnya sesuai dengan peraturan perundang-undangan yang berlaku di wilayah Negara Kesatuan Republik Indonesia.

Singaraja, 17 Juli 2020

Yang memberi pernyataan



Ida Bagus Nyoman Wijana Manuaba

PRAKATA

Puji syukur penulis panjatkan kehadapan Tuhan Yang Maha Esa, Ida Sang Hyang Widhi Wasa, karena berkat rahmat-Nya penulisan tesis ini dapat terselesaikan. Tesis penelitian dengan judul “*Analisis Sentimen Data Provider Layanan Internet Pada Twitter Menggunakan Support Vector Machine Dengan Penambahan Algoritma Levenshtein Distance*” ini disusun guna memenuhi persyaratan mencapai gelar Magister Komputer dari Program Pascasarjana Universitas Pendidikan Ganesha dan dengan harapan dapat menjadi pedoman dalam pelaksanaan penelitian yang diusulkan.

Dalam menyelesaikan tesis ini, penulis banyak mendapat bantuan baik berupa moral maupun material dari berbagai pihak. Untuk itu dalam kesempatan ini, penulis mengucapkan terima kasih kepada:

1. Dr. Gede Indrawan, S.T., M.T. selaku Kepala Program Magister Ilmu Komputer Undiksha sekaligus Pembimbing II yang telah memberikan bimbingan, arahan, petunjuk dan motivasi penulis baik selama perkuliahan maupun dalam penyusunan tesis ini.
2. Dr. Gede Rasben Dantes, S.T, M.T.I selaku Pembimbing I yang telah memberikan bimbingan, arahan, petunjuk dan motivasi penulis dalam penyelesaian tesis ini.
3. Seluruh staf dosen di lingkungan Program Magister Ilmu Komputer Universitas Pendidikan Ganesha, seluruh keluarga dan rekan-rekan mahasiswa yang tidak dapat yang telah banyak memberikan bantuan dan motivasi kepada penulis dalam penyusunan tesis ini.

Penulis menyadari bahwa penulisan tesis penelitian ini masih jauh dari kesempurnaan, untuk itu sangat diharapkan kritik maupun saran yang bersifat membangun, demi penyempurnaan penelitian ini.

Singaraja, 17 Juli 2020

Penulis

DAFTAR ISI

PRAKATA	ii
ABSTRAK	ii
ABSTRACT	iii
DAFTAR ISI	iv
DAFTAR TABEL	vii
DAFTAR GAMBAR	viii
BAB I PENDAHULUAN	1
1.1 Latar Belakang Penelitian	1
1.2 Rumusan Masalah	4
1.3 Batasan Masalah	4
1.4 Tujuan Penelitian	5
1.5 Manfaat Penelitian	5
BAB II KAJIAN PUSTAKA	6
2.1 Twitter	6
2.2 Analisis Sentimen	7
2.3 <i>Text Preprocessing</i>	8
2.4 <i>Feature Extraction</i>	9
2.5 Algoritma Levenshtein Distance	11
2.6 <i>SVM</i>	13
2.7 Kajian Hasil Penelitian yang Relevan	15
BAB III METODOLOGI	18
3.1 Prosedur Penelitian	18
3.1.1 Akusisi Data	19
3.1.2 <i>Feature Extraction</i>	24
3.1.3 <i>SVM Training</i>	24
3.1.4 <i>SVM Testing</i>	25
3.1.5 Hasil Klasifikasi <i>SVM</i>	25

3.1.6	Pengujian.....	25
3.1.7	Analisis Hasil	26
3.2	Implementasi	26
BAB IV	HASIL DAN PEMBAHASAN	27
4.1	Pengumpulan Data Twitter.....	27
4.2	<i>Text Preprocessing</i> Data Komentar Twitter.....	31
4.2.1	<i>Convert Emoticon</i>	32
4.2.2	<i>Cleansing</i>	33
4.2.3	<i>Case Folding</i>	34
4.2.4	<i>Stemming</i>	35
4.2.5	<i>Stopword Removal</i>	36
4.2.6	<i>Tokenizing</i>	38
4.2.7	Penerapan Algoritma Leveinsthein Distance	38
4.2.8	<i>Stopword Removal Lanjutan</i>	41
4.2.9	<i>Convert Negation</i>	42
4.3	<i>Feature Extraction</i>	42
4.3.1	<i>Term Frequency (TF)</i>	43
4.3.2	<i>Term Frequency - Invers Document Frequency (TF-IDF)</i>	43
4.4	<i>SVM</i>	44
4.3	Pengujian <i>SVM</i>	45
4.3.1	Pengujian <i>SVM</i> Tanpa Menggunakan <i>Algoritma Levensthein Distance</i> pada <i>Text Preprocessing</i>	46
4.3.2	Pengujian <i>SVM</i> dengan Menggunakan <i>Algoritma Levensthein Distance</i> pada <i>Text preprocessing</i>	47
4.4	Analisis Hasil.....	48
BAB V	PENUTUP	51
5.1	Rangkuman.....	51
5.2	Simpulan.....	52

5.3	Saran	53
DAFTAR PUSTAKA		54



DAFTAR TABEL

Tabel 2. 1 Tabel Matriks Perhitungan <i>Edit Distance</i>	12
Tabel 3. 1 Data Komentar yang Sama	20
Tabel 3. 2 <i>Confusion Matrix</i>	25
Tabel 4. 1 <i>Key API Twitter</i>	27
Tabel 4. 2 Hasil Pengumpulan Data Menggunakan <i>API Twitter</i>	28
Tabel 4. 3 Hasil Pemberian Label <i>Judges</i>	29
Tabel 4. 4 Jumlah Hasil Label Setiap Kelas Data	29
Tabel 4. 5 Jumlah Data <i>Training</i> dan Data <i>Testing</i>	30
Tabel 4. 6 Contoh Komentar Pada Twitter	31
Tabel 4. 7 Daftar Emoticon.....	32
Tabel 4. 8 Hasil Proses <i>Cleansing</i>	33
Tabel 4. 9 Hasil Tahapan <i>Case Folding</i>	34
Tabel 4. 10 Hasil Tahapan <i>Stemming</i>	35
Tabel 4. 11 Kamus <i>Stopword</i>	37
Tabel 4. 12 Hasil Tahapan <i>Stopword Removal</i>	37
Tabel 4. 13 Hasil Tahapan <i>Tokenizing</i>	38
Tabel 4. 14 Kamus Kata Dasar Bahasa Indonesia.....	39
Tabel 4. 15 Daftar Kata <i>Defined Word</i>	39
Tabel 4. 16 Hasil Bobot <i>Algoritma Levensthein Distance</i>	40
Tabel 4. 17 Daftar Kata Negasi	42
Tabel 4. 18 <i>Term Frequency</i>	43
Tabel 4. 19 Tabel Hasil <i>TF-IDF</i>	43
Tabel 4. 20 Perbandingan Hasil Klasifikasi Data <i>Testing</i>	44
Tabel 4. 21 <i>Confusion Matrix</i> Pengujian <i>SVM</i> Tanpa Menggunakan <i>Algoritma Levensthein Distance</i> pada <i>Text Preprocessing</i>	46
Tabel 4. 22 <i>Confusion Matrix</i> Pengujian <i>SVM</i> Dengan Menggunakan <i>Algoritma Levensthein Distance</i> pada <i>Text Preprocessing</i>	47

DAFTAR GAMBAR

Gambar 2. 1 <i>Hyperpalne</i> Memisahkan Data Kelas Positif dan Kelas Negatif	14
Gambar 3. 1 Prosedur Penelitian	18
Gambar 3. 2 Tahapan Akusisi Data	19
Gambar 3. 3 Tahapan <i>Text Preprocessing</i> pada Data Twitter	22
Gambar 3. 4 Proses Penerapan Algoritma <i>Levensthein Distance</i>	23
Gambar 4. 1 Tampilan Data Komentar <i>Twitter</i> Pada Sistem	31
Gambar 4. 2 Grafik Perbandingan Dua Hasil Pengujian	48
Gambar 4. 3 Perbedaan Waktu Proses pada <i>Text Preprocessing</i>	50

