

**KLASIFIKASI TEKS BERITA BERBAHASA INDONESIA
DENGAN ALGORITMA NAIVE BAYES MELALUI
KOMBINASI LAVENSTEIN DISTANCE DAN TERJEMAHAN
KATA BAHASA INGGRIS-INDONESIA**

TESIS



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



**KLASIFIKASI TEKS BERITA BERBAHASA INDONESIA
DENGAN ALGORITMA NAIVE BAYES MELALUI
KOMBINASI LAVENSTEIN DISTANCE DAN TERJEMAHAN
KATA BAHASA INGGRIS-INDONESIA**

TESIS



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

Tesis oleh I Gede Wira Kusuma Jaya ini telah diperiksa dan disetujui untuk mengikuti Ujian Tesis.

Singaraja, 10 Mei 2020

Pembimbing I,



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

NIP. 19760102200312 1 001

Pembimbing II,



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

NIP. 19760102 200312 1 001

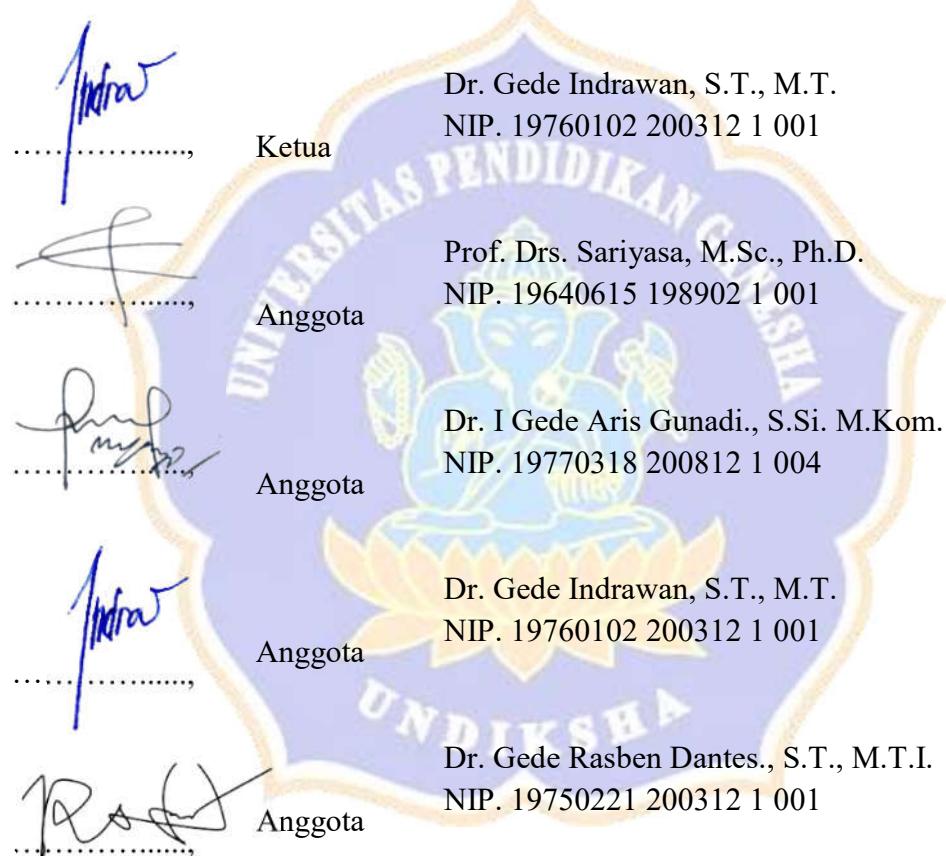


Tesis oleh I Gede Wira Kusuma Jaya ini telah berhasil dipertahankan di depan tim penguji dan dinyatakan diterima sebagai sebagian persyaratan untuk memperoleh gelar Magister Komputer di Program Studi Ilmu Komputer, Program Pascasarjana, Universitas Pendidikan Ganesha.

Disetujui pada tanggal : 3 Juni 2020

oleh

Tim Penguji



Mengetahui Direktur
Program Pascasarjana UNDIKSHA,



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 hasil karya orang lain telah dituliskan sumbernya secara jelas dan sesuai dengan norma, kaidah, serta etika akademis.

Apabila dikemudian 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, 6 Juni 2020

Yang membuat pernyataan



I Gede Wira Kusuma Jaya

PRAKATA

Puji dan syukur penulis panjatkan kehadapan Tuhan Yang Maha Esa atas anugrah-Nya, sehingga tesis yang berjudul “” dapat diselesaikan sesuai dengan yang direncanakan.

Tesis ini ditulis untuk memenuhi sebagian persyaratan dalam menyelesaikan studi di Program Studi Ilmu Komputer, Program Pascasarjana, Universitas Pendidikan Ganesha,. Pada lembar prakata tesis ini, ijinkan penulis menyampaikan terimakasih dan penghargaan yang sebesar-besarnya kepada :

1. Dr. Gede Rasben Dantes., S.T., M.T.I. sebagai pembimbing I dan Ketua Program Studi Ilmu Komputer yang dengan sabar membimbing, mengarahkan, dan memberikan motivasi yang demikian bermakna, sehingga penulis mampu melewati berbagai kerikil dalam perjalanan studi dan penyelesaian tesis ini.
2. Dr. Gede Indrawan., S.T., M. T., sebagai pembimbing II yang telah membimbing, mengarahkan, dan memberikan motivasi selama penyusunan tesis, sehingga tesis ini dapat terwujud dengan baik sesuai harapan.
3. Rektor Universitas Pendidikan Ganesha, yang telah memfasilitasi berbagai kepentingan studi, selama penulis menempuh perkuliahan di Program Pascasarjana Undiksha.
4. Direktur Program Pascasarjana Undiksha dan staf, yang telah banyak membantu selama penulis mengikuti dan menyelesaikan penulisan tesis ini.
5. Staf dosen pengajar di Program Studi Ilmu Komputer yang telah banyak membantu dan memotivasi penulis selama perjalanan studi dan penyusunan tesis ini.
6. Rekan-rekan seangkatan di Program Studi Ilmu Komputer yang dengan karakternya masing-masing telah banyak berkontribusi membentuk kemandirian penulis selama menjalani studi dan menyelesaikan tesis ini.
7. Bapak I Nengah Artika dan Ibu Ni Kadek Ringsun selaku orang tua penulis, serta adik dan keluarga yang telah banyak membantu secara material dan moral selama perjalanan studi yang penulis lakoni di Program Studi Ilmu Komputer

Semoga semua bantuan yang telah mereka taburkan dalam perjalanan studi penulis, terhargakan dengan sepantasnya oleh Tuhan Yang Maha Esa, sehingga mereka diberi jalan, rejeki, dan keharmonian dalam menjalani setiap langkah kehidupan.

Penulis menyadari sepenuhnya bahwa tesis ini masih jauh dari sempurna, untuk itu semua jenis saran, kritik dan masukan yang bersifat membangun sangat penulis harapkan. Akhir kata, semoga tesis ini dapat memberikan manfaat dan memberikan wawasan tambahan bagi para pembaca.



DAFTAR ISI

PRAKATA	iv
ABSTRAK	vi
<i>ABSTRACT</i>	vii
Daftar Isi.....	viii
Daftar Gambar.....	ix
Daftar Tabel	x
BAB I Pendahuluan	1
1.1 Latar Belakang	1
1.3 Batasan Masalah.....	4
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Hasil Penelitian	5
BAB II Kajian Pustaka.....	6
2.1 Landasan Teori	6
2.1.1 Preproses Teks	6
2.2.1 Algoritma Naive Bayes	9
2.2 Levenhstein Distance	15
2.3 Penelitian Terkait	17
BAB III METODE PENELITIAN.....	22
3.1 Tahapan Penelitian	22
BAB IV	41
Hasil dan Pembahasan.....	41
4.1 Hasil Percobaan.....	41
4.2 Perbandingan Hasil.....	53
BAB V Penutup	56
5.1 Rangkuman.....	56
5.2 Simpulan.....	57
5.3 Saran	58
Daftar Pustaka	59
Lampiran	61

DAFTAR GAMBAR

Gambar 2.1 Text Preprocessing	17
Gambar 2.2 Metodologi	18
Gambar 3.1 Tahapan Penelitian	22
Gambar 4.1 Perbandingan Akurasi Klasifikasi Percobaan Ke-1	42
Gambar 4.2 Perbandingan Akurasi Klasifikasi Percobaan Ke-2	45
Gambar 4.3 Perbandingan Akurasi Klasifikasi Percobaan Ke-3	47
Gambar 4.4 Perbandingan Akurasi Klasifikasi Percobaan Ke-4	49
Gambar 4.5 Perbandingan Akurasi Klasifikasi Percobaan Ke-5	52
Gambar 4.6 Grafik Perbandingan Rata-rata Akurasi Setiap Percobaan	54



DAFTAR TABEL

Tabel 2. 1 Probabilitas.....	10
Tabel 2. 4 Perbandingan Akurasi Setiap Kombinasi Algoritma	19
Tabel 3.1 Frekuensi Kemunculan Kata.....	29
Tabel 3. 2 Naive Bayes, Frekuensi Kemunculan Kata.....	30
Tabel 3.3 Confusion Matrix	39
Tabel 4.1 Confusion Matrix Naive Bayes Percobaan Ke-1	41
Tabel 4.2 Confusion Matrix Modifikasi Naive Bayes Percobaan Ke-1.....	42
Tabel 4.3 Precision Klasifikasi Percobaan Ke-1	43
Tabel 4.4 Recall Klasifikasi Percobaan Ke-1	43
Tabel 4.5 Confusion Matrix Naive Bayes Percobaan Ke-2	44
Tabel 4.6 Confusion Matrix Modifikasi Naive Bayes Percobaan Ke-2.....	44
Tabel 4.7 Precision Klasifikasi Percobaan Ke-2.....	45
Tabel 4.8 Recall Klasifikasi Percobaan Ke-2	45
Tabel 4.9 Confusion Matrix Naive Bayes Percobaan Ke-3	46
Tabel 4.10 Confusion Matrix Modifikasi Naive Bayes Percobaan Ke-3.....	46
Tabel 4.11 Precision Klasifikasi Percobaan Ke-3	47
Tabel 4.12 Recall Klasifikasi Percobaan Ke-3	48
Tabel 4.13 Confusion Matrix Naive Bayes Percobaan Ke-4	48
Tabel 4.14 Confusion Matrix Modifikasi Naive Bayes Percobaan Ke-4.....	49
Tabel 4.15 Precision Klasifikasi Percobaan Ke-4	50
Tabel 4.16 Recall Klasifikasi Percobaan Ke-4	50
Tabel 4.17 Confusion Matrix Naive Bayes Percobaan Ke-5	51
Tabel 4.18 Confusion Matrix Modifikasi Naive Bayes Percobaan Ke-5.....	51
Tabel 4.19 Precision Klasifikasi Percobaan Ke-5	52
Tabel 4.20 Recall Klasifikasi Percobaan Ke-5	52
Tabel 4.21 Rata-rata Perbandingan Akurasi Klasifikasi	53
Tabel 4.22 Perbandingan Precision dan Recall.....	54