

**ANALISIS SENTIMEN KOMENTAR YOUTUBE
TERHADAP KEBIJAKAN PAJAK DAN INSENTIF
KENDARAAN LISTRIK DI INDONESIA
MENGUNAKAN NAÏVE BAYES DAN
K-NEAREST NEIGHBOR**

TESIS



**OLEH:
I NYOMAN WIRYA YASA
2129101058**

**PROGRAM STUDI ILMU KOMPUTER (S2)
PROGRAM PASCASARJANA
UNIVERSITAS PENDIDIKAN GANESHA**



- UU ITE No. 11 Tahun 2008 Pasal 5 Ayat 1 "Informasi Elektronik dan/atau hasil cetaknya merupakan alat bukti hukum yang sah"
- Dokumen ini telah ditandatangani secara elektronik menggunakan sertifikat elektronik yang diterbitkan BSR - BSSN, validitas dokumen elektronik ini bisa dicek menggunakan aplikasi mobile VeryDS oleh BSR
- Cetakan dokumen ini merupakan salinan dari file dokumen bertandatangan elektronik yang keabsahannya dapat diakses melalui scan QRCode yang terdapat pada sertifikat ini.



- UU ITE No. 11 Tahun 2008 Pasal 5 Ayat 1 "Informasi Elektronik dan/atau hasil cetaknya merupakan alat bukti hukum yang sah"
- Dokumen ini telah ditandatangani secara elektronik menggunakan sertifikat elektronik yang diterbitkan BSR - BSSN, validitas dokumen elektronik ini bisa dicek menggunakan aplikasi mobile VeryDS oleh BSR
- Cetakan dokumen ini merupakan salinan dari file dokumen bertandatangan elektronik yang keabsahannya dapat diakses melalui scan QRCode yang terdapat pada sertifikat ini.



Balai
Sertifikasi
Elektronik

TESIS

DIAJUKAN UNTUK MELENGKAPI TUGAS DAN MEMENUHI SYARAT-SYARAT UNTUK MENCAPAI GELAR MAGISTER KOMPUTER

	
Menyetujui	
Pembimbing I	Dr. Ir. I Made Gede Sunarya, S.Kom., M.Cs. NIP.198307252008011008
Pembimbing II	Dr. Putu Hendra Suputra, S.Kom., M.Cs. NIP.198212222006041001



- UU ITE No. 11 Tahun 2008 Pasal 5 Ayat 1 "Informasi Elektronik dan/atau hasil cetaknya merupakan alat bukti hukum yang sah"
- Dokumen ini telah ditandatangani secara elektronik menggunakan sertifikat elektronik yang diterbitkan BSrE - BSSN, validitas dokumen elektronik ini bisa dicek menggunakan aplikasi mobile VeryDS oleh BSrE
- Cetakan dokumen ini merupakan salinan dari file dokumen bertandatangan elektronik yang keabsahannya dapat diakses melalui scan QRCode yang terdapat pada sertifikat ini.

Tesis oleh I Nyoman Wirya Yasa telah dipertahankan di depan tim penguji dan dinyatakan diterima sebagai salah satu persyaratan untuk memperoleh gelar Magister Komputer di Ilmu Komputer (S2), Program Pascasarjana, Universitas Pendidikan Ganesha
Pada tanggal 05 Agustus 2026

Dewan Penguji

Ketua	Dr. Ir. I Made Gede Sunarya, S.Kom., M.Cs. NIP.198307252008011008
Anggota	Dr. Putu Hendra Suputra, S.Kom., M.Cs. NIP.198212222006041001
Anggota	Prof. Dr. Gede Indrawan, S.T., M.T. NIP.197601022003121001
Anggota	Dr. I Gede Aris Gunadi, S.Si. M.Kom. NIP.197703182008121004

Mengetahui Direktur Program Pascasarjana
Universitas Pendidikan Ganesha,



Prof. Dr. I Nyoman Jampel, M.Pd.

NIP.195910101986031003



- UU ITE No. 11 Tahun 2008 Pasal 5 Ayat 1 "Informasi Elektronik dan/atau hasil cetaknya merupakan alat bukti hukum yang sah"
- Dokumen ini telah ditandatangani secara elektronik menggunakan sertifikat elektronik yang diterbitkan BSrE - BSSN, validitas dokumen elektronik ini bisa dicek menggunakan aplikasi mobile VeryDS oleh BSrE
- Cetakan dokumen ini merupakan salinan dari file dokumen bertandatangan elektronik yang keabsahannya dapat diakses melalui scan QRCode yang terdapat pada sertifikat ini.

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 di kemudian hari ditemukan seluruh atau sebagian tesis ini bukan hasil karya saya sendiri atau adanya plagiat dalam bagianbagian 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.



KATA PENGANTAR

Puji dan syukur penulis panjatkan ke hadapan Tuhan Yang Maha Esa atas anugerah-Nya, sehingga tesis yang berjudul “Analisis Sentimen Komentar Youtube Terhadap Kebijakan Pajak dan Insentif Kendaraan Listrik di Indonesia Menggunakan *Naïve Bayes* dan *K-Nearest Neighbor*” dapat terselesaikan sesuai dengan yang direncanakan.

Selama penyusunan tesis ini, penulis menemui beberapa kendala dan kesulitan, namun berkat dukungan dan bantuan dari berbagai pihak, penulis dapat mengatasi hambatan dan kesulitan tersebut, sehingga pada kesempatan ini penulis ingin menyampaikan rasa terimakasih yang sebesar-besarnya kepada:

1. Bapak Rektor Universitas Pendidikan Ganesha, yang telah memberikan dukungan secara moril dan memfasilitasi berbagai kepentingan akademik, selama Penulis menempuh studi di Program Pascasarjana Universitas Pendidikan Ganesha.
2. Bapak Direktur Program Pascasarjana Universitas Pendidikan Ganesha, yang telah banyak memberikan bimbingan dan arahan serta petunjuk hingga terselesaikannya tesis ini.
3. Koordinator Program Studi Ilmu Komputer Universitas Pendidikan Ganesha yang telah banyak membantu dan menyemangati penulis selama menjalani pendidikan dan penyusunan tesis ini.
4. Bapak Dr. Ir. I Made Gede Sunarya, S.Kom., M.Cs., sebagai Dosen Pembimbing I yang telah memberikan bimbingan, arahan dan petunjuk yang bermanfaat serta memacu semangat Penulis untuk menyelesaikan tesis ini sesuai yang diharapkan.
5. Bapak Dr. Putu Hendra Suputra, S.Kom., M.Cs., selaku Dosen Pembimbing II yang dengan sabar memberikan bimbingan, arahan, dan memberikan motivasi yang demikian bermakna, sehingga penulis mampu mengatasi berbagai kendala dalam perjalanan studi dan penyelesaian tesis ini.
6. Bapak Prof. Dr. Gede Indrawan, M.T., selaku Dosen Penguji I yang dengan sabar memberikan bimbingan, arahan, dan memberikan motivasi yang

demikian bermakna, sehingga penulis mampu mengatasi berbagai kendala dalam perjalanan studi dan penyelesaian tesis ini.

7. Bapak Dr. I Gede Aris Gunadi, S.Si., M.Kom. selaku Dosen Penguji II sabar memberikan bimbingan, arahan, dan memberikan motivasi yang demikian bermakna, sehingga penulis mampu mengatasi berbagai kendala dalam perjalanan studi dan penyelesaian tesis ini.
8. Bapak dan Ibu dosen dan Staf, yang telah memberikan pelajaran kepada Penulis selama mengikuti studi di Program Studi Ilmu Komputer Universitas Pendidikan Ganesha.
9. Segenap keluarga tercinta, Bapak, Ibu, dan Saudara dan Istri tercinta yang telah banyak membantu memberikan semangat dan motivasi serta dukungan moril dan materiil dalam penulisan tesis ini.
10. Semua rekan, sahabat dan pihak-pihak yang tidak dapat Penulis sebutkan satu per satu atas bantuan secara langsung maupun tidak langsung dalam penulisan tesis ini.

Akhir kata Penulis mengucapkan terima kasih kepada semua pihak yang telah berkontribusi dalam penulisan tesis ini dan Penulis berharap semoga tesis ini dapat bermanfaat bagi kita semua dan dunia pendidikan di masa yang akan datang.

Singaraja, 5 Agustus 2026

Penulis

DAFTAR ISI

	Halaman
LEMBAR PERSETUJUAN PEMBIMBING	iv
LEMBAR PERSETUJUAN PENGUJI.....	v
LEMBAR PERNYATAAN.....	vi
KATA PENGANTAR.....	vii
ABSTRAK	ix
ABSTRACT.....	x
DAFTAR ISI.....	xi
DAFTAR GAMBAR.....	xiii
DAFTAR TABEL	xiv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah	4
1.3 Batasan Penelitian.....	5
1.4 Rumusan Masalah.....	5
1.5 Tujuan Penelitian	5
1.6 Manfaat Penelitian	6
BAB II LANDASAN TEORI.....	8
2.1 Kajian Pustaka	8
2.1.1 Analisis sentimen.....	8
2.1.2 <i>Preprocessing</i>	9
2.1.3 TF-IDF.....	10
2.1.4 <i>Naïve Bayes Classifier</i> (NBC).....	11
2.1.5 <i>K-Nearest Neighbors</i> (K-NN)	13
2.1.6 <i>Confusion Matrix</i>	15
2.2 Kajian Penelitian yang Relevan	18
2.3 Kerangka Berpikir.....	23
BAB III METODE PENELITIAN.....	24
3.1 Alur Penelitian	24
3.2 Studi Pustaka.....	24
3.3 Pengumpulan Data	25
3.4 Pelabelan Data	26
3.5 <i>Preprocessing</i> Data.....	28

3.6	Proses Pembobotan Kata (Term TF-IDF).....	32
3.7	Analisis Data Klasifikasi.....	37
3.7.1	Klasifikasi <i>Naïve Bayes</i>	38
3.7.2	<i>K-Nearest Neighbor</i>	41
3.8	Evaluasi dan Perbandingan <i>Performance</i>	43
BAB IV	HASIL DAN PEMBAHASAN	45
4.1	Hasil Pengumpulan Data.....	45
4.2	Hasil Data <i>Preprocessing</i>	46
4.3	Pembobotan Kata (TF-IDF).....	51
4.4	Skenario Pengujian	52
4.5	Penanganan Ketidakseimbangan Data.....	54
4.6	Hasil Implementasi Klasifikasi.....	55
4.6.1	Hasil klasifikasi <i>Naïve Bayes</i>	55
4.6.2	Hasil klasifikasi <i>K-Nearest Neighbor</i>	58
4.7	Perbandingan Hasil Metode <i>Naïve Bayes Classifier</i> dan <i>K-Nearest Neighbor</i>	59
4.7.1	Akurasi, Presisi, <i>Recall</i> , dan <i>F1-Score</i>	60
4.7.2	<i>Receiver Operating Characteristic – Area Under the Curve</i> (ROC-AUC)	62
BAB V	PENUTUP.....	67
5.1	Rangkuman	67
5.2	Kesimpulan	68
5.3	Keterbatasan.....	73
5.4	Saran	74
DAFTAR PUSTAKA		77
LAMPIRAN		79

DAFTAR GAMBAR

	Halaman
Gambar 2.1 Algoritma K-Nearest Neighbor	14
Gambar 3.1 Alur Penelitian.....	24
Gambar 3.2 Perbandingan Komentar Negatif dan Positif.....	27
Gambar 4.1 Data Awal Komentar.....	47
Gambar 4.2 Data Duplikat.....	47
Gambar 4.3 Hasil Proses <i>Cleaning</i>	48
Gambar 4.4 Hasil Proses <i>Case Folding</i>	48
Gambar 4.5 Hasil Proses Normalisasi.....	49
Gambar 4.6 Hasil Proses <i>Tokenizing</i>	49
Gambar 4.7 Hasil Proses <i>Stopword Removal</i>	50
Gambar 4.8 Hasil Proses <i>Stemming</i>	50
Gambar 4.9 Visualisasi <i>Word Cloud</i>	51
Gambar 4.10 Visualisasi <i>Word Frequency</i>	51
Gambar 4.11 Hasil Pembobotan Kata (TF-IDF)	52
Gambar 4.12 Pembagian Dataset dengan 80% Data Latih dan 20% Data Uji..	53
Gambar 4.13 Penyeimbangan Data dengan SMOTE.....	55
Gambar 4.14 Proses Melatih Model <i>Naïve Bayes</i>	56
Gambar 4.15 Hasil Prediksi dari Model <i>Naïve Bayes</i>	56
Gambar 4.16 Hasil Proses Evaluasi Klasifikasi <i>Naive Bayes</i>	56
Gambar 4.17 Grafik Perbandingan <i>Accuracy, Precision, Recall, dan F1-Score</i> (Data Tidak Seimbang/Tanpa SMOTE).....	60
Gambar 4.18 Grafik Perbandingan <i>Accuracy, Precision, Recall, dan F1-Score</i> (Data Seimbang/dengan SMOTE).....	61
Gambar 4.19 Kurva ROC NBC dan K-NN (Data Tidak Seimbang)	63
Gambar 4.20 Kurva ROC NBC dan K-NN (Data Seimbang).....	64
Gambar 4.21 Grafik Perbandingan ROC-AUC.....	65

DAFTAR TABEL

	Halaman
Tabel 2.1 <i>Confusion matrix</i>	16
Tabel 2.2 Penelitian Terkait	18
Tabel 3.1 Daftar Video Sumber Data	25
Tabel 3.2 Komentar Isu Kebijakan Pajak dan Insentif pada Kendaraan Listrik	26
Tabel 3.3 Komentar dan Label dari Isu Kebijakan Pajak dan Insentif pada Kendaraan Listrik	28
Tabel 3.4 Data Duplikat pada Kolom Komentar	28
Tabel 3.5 Hasil Clening Elemen Teks yang tidak Relevan	29
Tabel 3.6 Hasil Proses <i>Case Folding</i>	30
Tabel 3.7 Hasil Proses Normalisasi	30
Tabel 3.8 Hasil Proses <i>Tokenizing</i>	31
Tabel 3.9 Hasil Proses <i>Stopword Removal</i>	32
Tabel 3.10 Hasil Proses <i>Stemming</i>	32
Tabel 3.11 Frekuensi Kemunculan Kata	33
Tabel 3.12 Sampel Dokumen dan Hasil <i>Preprocessing</i>	34
Tabel 3.13 Proses Perhitungan TF	34
Tabel 3.14 Proses Perhitungan IDF	35
Tabel 3.15 Hasil Perhitungan TF-IDF	36
Tabel 3.16 Data Latih dan Data Uji	38
Tabel 3.17 Probabilitas Setiap Kelas	39
Tabel 3.18 Probabilitas Setiap Kata Terhadap Setiap Kelas	39
Tabel 3.19 Perhitungan <i>Likelihood</i>	41
Tabel 3.20 Perhitungan <i>Euclidean Distance</i>	42
Tabel 3.21 Pengurutan Nilai <i>Euclidean Distance</i>	43
Tabel 4.1 Contoh Komentar dengan Sentimen Positif, Negatif, dan Netral	46
Tabel 4.2 Hasil Klasifikasi <i>Metode Naïve Bayes</i>	57
Tabel 4.3 Hasil Evaluasi Klasifikasi <i>K-Nearest Neighbor</i>	58