

Rini Indriati, M.Kom.

Analisis Sentimen

UNDANG-UNDANG CIPTA KERJA

PADA TWITTER



Analisis Sentimen

UNDANG-UNDANG CIPTA KERJA

PADA TWITTER

Rini Indriati, M.Kom.

 Penerbit
litrus.

**ANALISIS SENTIMEN UNDANG-UNDANG CIPTA KERJA
PADA TWITTER**

Ditulis oleh :
Rini Indriati, M.Kom.

Diterbitkan, dicetak, dan didistribusikan oleh
PT. Literasi Nusantara Abadi Grup
Perumahan Puncak Joyo Agung Residence Kav. B11 Merjosari
Kecamatan Lowokwaru Kota Malang 65144
Telp : +6285887254603, +6285841411519
Email: literasinusantaraofficial@gmail.com
Web: www.penerbitlitnus.co.id
Anggota IKAPI No. 340/JTI/2022



Hak Cipta dilindungi oleh undang-undang. Dilarang mengutip atau memperbanyak baik sebagian ataupun keseluruhan isi buku dengan cara apa pun tanpa izin tertulis dari penerbit.

Cetakan I, Oktober 2024

Perancang sampul: Hasanuddin
Penata letak: Noufal Fahriza

ISBN : 978-623-519-481-3
viii + 116 hlm. ; 15,5x23 cm.

©Oktober 2024



KATA PENGANTAR

Puji syukur dipanjatkan pada Tuhan yang Maha Pengasih Dan Penyayang, atas RahmatNya penyusunan buku monograf ini dapat terselesaikan.

Buku ini merupakan hasil kajian yang berfokus pada bidang ilmu Analisis Sentimen (*sentiment analysis*), suatu metode yang semakin berkembang pesat dalam era digital. Analisis Sentimen mengacu pada proses analisis teks digital yang bertujuan untuk menentukan apakah suatu teks memiliki muatan emosional positif atau negatif. Metode ini sangat berguna dalam memahami pandangan dan persepsi publik terhadap berbagai topik, salah satunya adalah kebijakan pemerintah.

Dalam monograf ini, kami menggunakan platform media sosial Twitter sebagai sumber pengambilan teks digital. Twitter dipilih karena popularitasnya yang tinggi di kalangan masyarakat Indonesia. Sebagai salah satu aplikasi media sosial terbesar, Twitter memungkinkan pengguna untuk mengirimkan dan menerima pesan secara real-time, yang menjadikannya sebagai ruang publik yang dinamis dan efektif dalam mengukur opini serta persepsi masyarakat. Melalui karakteristik ini, Twitter menjadi media ideal untuk memonitor sentimen publik terhadap kebijakan besar seperti Undang-Undang Cipta Kerja.

Proses kajian dalam buku ini mencakup pemanfaatan teknik analisis sentimen untuk menelaah bagaimana masyarakat bereaksi terhadap UU Cipta Kerja melalui platform Twitter. UU Cipta Kerja merupakan kebijakan yang memunculkan diskusi luas dan beragam

pandangan di berbagai kalangan. Reaksi terhadap kebijakan ini sangat bervariasi, dari yang mendukung hingga yang menentang, dan Twitter menjadi salah satu media di mana opini-opini ini terpapar dengan jelas. Dengan menganalisis teks-teks yang berkaitan dengan UU Cipta Kerja, kajian ini memberikan gambaran mengenai pandangan masyarakat terhadap kebijakan tersebut, baik dari sudut pandang positif maupun negatif.

Harapan kami, buku ini tidak hanya bermanfaat sebagai karya ilmiah, tetapi juga sebagai rujukan yang berguna bagi berbagai kalangan, terutama mahasiswa dan dosen dalam bidang teknologi informasi. Bagi mereka, buku ini akan menjadi contoh konkrit penerapan analisis sentimen dalam studi ilmiah berbasis teknologi. Selain itu, dunia usaha dan industri juga dapat mengambil manfaat dari buku ini, mengingat analisis sentimen kini menjadi salah satu alat penting dalam memahami tren pasar dan persepsi konsumen. Melalui studi kasus yang didalami dalam buku ini, para pelaku industri dapat belajar bagaimana sentimen publik dapat mempengaruhi pandangan dan keputusan bisnis terhadap kebijakan tertentu.

Tak hanya bagi akademisi dan pelaku bisnis, masyarakat umum pun diharapkan dapat memperoleh wawasan baru mengenai cara pandang publik terhadap kebijakan-kebijakan yang memengaruhi kehidupan sehari-hari. Dalam era informasi yang serba cepat ini, opini publik yang tersebar melalui media sosial memiliki kekuatan yang signifikan dalam membentuk persepsi kolektif, dan analisis sentimen menyediakan alat untuk memetakan dinamika tersebut secara lebih sistematis.

Buku ini juga diharapkan bisa menjadi panduan awal bagi para peneliti yang ingin mendalami lebih lanjut mengenai penerapan teknologi analisis sentimen dalam kajian sosial-politik. Melalui hasil kajian ini, kami berharap dapat mendorong studi-studi lanjutan yang lebih mendalam di masa depan, khususnya dalam menghubungkan antara dinamika opini publik di media sosial dengan dampaknya terhadap pengambilan keputusan kebijakan.

Terimakasih kepada semua pihak atas kerjasama dan dukungan atas penyelesaian buku ini. Penyusunan buku ini masih banyak kekurangan dan masih perlu penyempurnaan, oleh karena itu diharapkan ada kritik atau masukan agar buku ini menjadi lebih baik.

Kediri, 2024

Rini Indriati, M.Kom.



DAFTAR ISI

Kata Pengantar	iii	Estimasi Parameter dan Model Kejadian	19
Daftar Isi.....	vii	Kelebihan dan Kekurangan <i>Naïve Bayes Classifier</i>	27
Daftar Gambar	ix		
Daftar Tabel	xi		

BAB I

PENDAHULUAN.....	1
------------------	---

BAB II

ANALISIS SENTIMEN..	7
---------------------	---

Pengertian Analisis Sentimen .	7
Cara Kerja Analisis Sentimen .	9
Metode Analisis Sentimen	9
Jenis Analisis Sentimen	11
Tantangan Analisis Sentimen	12

BAB III

NAÏVE BAYES CLASSIFIER	15
------------------------------	----

Pengantar <i>Naïve Bayes</i>	16
Model Probabilistik (<i>Probabilistic Model</i>)	17
Membangun Pengklasifikasi dari Model Probabilitas	19

BAB IV

TERM FREQUENCY-INVERS DOCUMENT FREQUENCY (TF-IDF).....	29
--	----

Pengertian TF-IDF	30
Term Frekuensi-Frekuensi Dokumen Terbalik (<i>Term Frequency-Inverse Document Frequency</i>).....	32
Pembenaran <i>idf</i>	33
Korelasi dengan Teori Informasi	34
Derivatif	36

BAB V

WEB SCRAPING.....	37
-------------------	----

Pengertian.....	37
Teknik <i>Web Scraping</i>	39
Metode untuk Mencegah <i>Web Scraping</i>	43

BAB VI

MEDIA SOSIAL TWITTER/X..... 45

- Penampilan dan Fitur
(*Appearance and Features*) 46
- Penggunaan (*Usage*)..... 51
- Teknologi..... 53
- Akun Pengguna (*User
Accounts*)..... 56
- Masyarakat (*Society*) 58

BAB VII

WORD CLOUD DAN UNDANG-UNDANG CIPTA KERJA61

- Pengertian..... 61
- Undang-undang Cipta Kerja
dalam Kacamata Media
Sosial..... 65

BAB VIII

MENELIK SENTIMEN UU CIPTA KERJA DI TWITTER..... 69

- Pengumpulan Data 69
- Pelabelan Manual 70
- Pemodelan 74
- Pengujian Pemodelan 75
- Evaluasi Performa 76
- Alur Kajian 78

BAB IX

HASIL ANALISIS SENTIMEN UU CIPTA KERJA81

- Penghimpunan Data..... 81
- Pelabelan Manual 82
- Pengujian Performa 84
- Evaluasi Performa 86
- Word Cloud* Sentimen
- Undang-undang Cipta Kerja. 88

BAB X

PENUTUP91

- Glosarium 93
- Indeks..... 105
- Daftar Pustaka..... 111
- Profil Penulis..... 115



DAFTAR GAMBAR

Gambar 3.1	Contoh Pengklasifikasi Naïve Bayes sebagai Jaringan Bayesian.....	16
Gambar 6.1	Konten <i>tweet</i> menurut <i>Pear Analytics</i> pada bulan Agustus 2009	52
Gambar 7.1	Visualisasi Data Teks.....	62
Gambar 7.2	Tren Dan Pola Medsos.....	62
Gambar 8.1.	<i>Confusion Matrix</i>	77
Gambar 8.2.	Alur kajian	79
Gambar 9.1.	Sampel <i>tweet</i> hasil <i>web scraping</i>	81
Gambar 9.2.	Data tiap kata yang telah diberi bobot	84
Gambar 9.3.	Hasil Klasifikasi	86
Gambar 9.4.	Hasil Evaluasi Naïve Bayes.....	87
Gambar 9.5.	Sentimen Undang-Undang Cipta Kerja.....	88
Gambar 9.6.	<i>Word Cloud</i> Sentimen Positif.....	89
Gambar 9.7.	<i>Word Cloud</i> Sentimen Negatif.....	90



DAFTAR TABEL

Tabel 9.1	Frekuensi Dokumen Terbalik (<i>Inverse Document Frequency</i>).....	32
Tabel 9.2.	<i>Term Frequency-Inverse Document Frequency</i>	33
Tabel 9.1.	Data yang telah diberi label	82
Tabel 9.2.	Hasil <i>Preprocessing</i>	83
Tabel 9.3.	Hasil Pengujian.....	85



I

PENDAHULUAN

Analisis sentimen (*sentiment analysis*) adalah teknik pemrosesan bahasa alami (*Natural Language Processing*) yang digunakan untuk menentukan apakah data bersifat positif, negatif, atau netral. Pada analisis sentimen atau penggalian opini, merupakan proses menganalisis teks dalam jumlah besar untuk menentukan apakah teks tersebut mengekspresikan sentimen positif, sentimen negatif, atau sentimen netral.

Analisis sentimen berfokus pada polaritas sebuah teks (positif, negatif, netral) tetapi juga melampaui polaritas untuk mendeteksi perasaan dan emosi tertentu (seperti marah, bahagia, sedih, kecewa dan lain sebagainya), urgensi (misal: mendesak, tidak mendesak), niat (tertarik atau tidak tertarik).

Analisis sentimen multibahasa melibatkan banyak pra-pemrosesan dan sumber daya. Sebagian besar sumber daya tersedia *online*, sementara sumber daya lainnya perlu dibuat namun harus mengetahui cara membuat kode untuk menggunakannya. Alternatif untuk dapat mendeteksi bahasa dalam teks secara otomatis dengan

mengklasifikasikan bahasa kemudian melatih model analisis sentimen khusus untuk mengklasifikasikan teks dalam bahasa yang dipilih.

Analisis sentimen diperlukan karena manusia mengekspresikan pikiran dan perasaannya dengan lebih terbuka. Analisis sentimen menjadi alat penting untuk memantau dan memahami sentimen pada semua jenis data.

Oleh karena itu, untuk mengklasifikasi bahasa yang tersedia secara *online* pada media sosial *Twitter* yang memuat sentimen Undang-Undang Cipta Kerja (*Omnibus Law*) menggunakan metode *Naïve Bayes Classifier* untuk menentukan sentimen positif, sentimen negatif.

Opini publik atau pendapat umum adalah pendapat kelompok masyarakat atau sintesis dari pendapat dan diperoleh dari diskusi sosial dari pihak-pihak yang memiliki kaitan kepentingan. Agregat dari sikap dan kepercayaan biasanya dianut oleh populasi orang dewasa. Pemilihan opini publik didasarkan kepada jumlah mayoritas yang efektif, dan bukan pada jumlah mayoritas secara keseluruhan. Opini publik menggunakan subjek berupa permasalahan baru yang mengandung pernyataan yang bersifat kontroversial. Sifat dari pernyataannya adalah memiliki suatu hal yang bertentangan, dan menjadi reaksi pertama atau sebuah gagasan baru.

Opini personal terdiri atas kegiatan verbal dan non-verbal yang menyajikan citra dan interpretasi individual tentang objek tertentu, biasanya dalam bentuk isu yang diperdebatkan orang.

Pernyataan dari suatu opini dapat dilakukan secara aktif maupun pasif. Selain itu, opini juga dapat diungkapkan melalui perkataan yang jelas maupun kiasan, serta dapat memiliki konotasi maupun persepsi pribadi.

Opini publik identik dengan pengertian kebebasan, keterbukaan dalam mengungkapkan ide-ide, pendapat, keinginan, keluhan, kritik yang membangun dan kebebasan di dalam penulisan. Dengan kata lain, opini publik merupakan efek dari kebebasan dalam mengungkapkan ide-ide dan pendapat. Untuk memahami opini seseorang dan publik,



ANALISIS SENTIMEN

Pengertian Analisis Sentimen

Analisis sentimen yang juga dikenal sebagai *opinion mining* atau emosi kecerdasan buatan adalah penggunaan pemrosesan bahasa alami (*Natural Language Processing*), analisis teks (*Text Analysis*), linguistik komputasional dan biometrik (*Computational Linguistic and Biometric*) untuk mengidentifikasi, mengekstrak, mengukur dan mempelajari keadaan efektif dan informasi subyektif secara sistematis.

Tugas dasar dalam analisis sentimen adalah mengklasifikasikan polaritas teks tertentu pada tingkat dokumen, kalimat, atau fitur/ aspek pada pendapat yang diungkapkan dalam dokumen, kalimat, atau fitur/aspek entitas bersifat positif, negatif atau netral. Klasifikasi sentimen tingkat lanjut yang melampaui polaritas misalnya, melihat pada kondisi emosional seperti kesenangan, kemarahan, rasa tidak suka, kesedihan, ketakutan dan sebagainya. Jadi, analisis sentimen adalah proses menganalisa teks digital untuk menentukan emosional pesan atau pendapat bersifat positif, negatif atau netral. Data teks dalam volume besar seperti email, percakapan, komentar media

sosial, dan ulasan dapat dipindai dengan alat analisis sentimen secara otomatis untuk menentukan sikap penulis terhadap suatu topik.

Analisis sentimen merupakan hal yang penting karena memberikan beberapa manfaat dalam hal memberi wawasan yang objektif, membangun produk dan layanan yang lebih baik, menganalisis dalam skala besar, hasil waktu nyata.

Alat analisis sentimen berbasis kecerdasan buatan (AI) mampu memberikan hasil yang konsisten dan objektif. Alat analisis sentimen secara akurat mampu mengurutkan dan mengklasifikasi teks untuk menangkap emosi secara objektif memberi wawasan yang objektif

Analisis sentimen membantu dalam meningkatkan produk dan layanan berdasarkan umpan balik pelanggan yang spesifik. Teknologi kecerdasan buatan (AI) mengidentifikasi objek atau entitas yang diasosiasikan pelanggan dengan sentiment negatif.

Mining informasi dari sejumlah besar data tidak yang terstruktur, seperti email, *chatbot*, survey, catatan manajemen hubungan pelanggan (CRM), umpan balik produk.

Perangkat lunak analisis sentiment untuk mempelajari apa yang dirasakan pelanggan tentang merek, produk, dan layanan perusahaan secara waktu nyata (*real-time*) dan mengambil tindakan berdasarkan hasil temuan. Konfigurasi perangkat lunak untuk mengirim peringatan saat sentimen negatif terdeteksi untuk kata kunci tertentu.

Beberapa kasus yang menggunakan analisis sentimen antara lain meningkatkan mutu layanan, yaitu mempersonalisasi respon berdasarkan percakapan pada *chatbox* dan dieskalasi ke staf dukungan; pemantauan merek dengan terus memantau penyebutan dan obrolan seputar merek di media sosial, forum, blog, artikel berita, dan ruang digital lainnya; Riset Pasar dengan mempelajari apa yang berhasil dan apa yang tidak berhasil, analisis komentar di situs ulasan online, jawaban survei, dan posting media sosial untuk mendapatkan wawasan yang lebih dalam tentang fitur produk tertentu; melacak performa promosi yaitu untuk memastikan promosi iklan menghasilkan



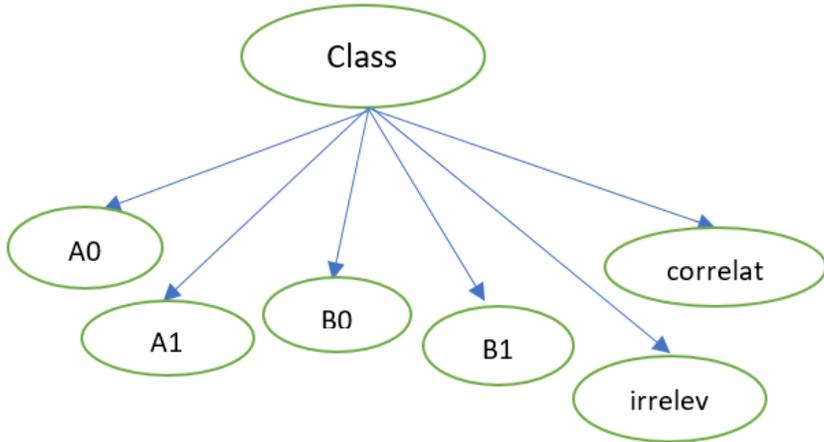
NAÏVE BAYES CLASSIFIER

Classifier *Naïve Bayes* adalah algoritma pembelajaran mesin terawasi (*supervised machine learning*) yang digunakan untuk tugas klasifikasi seperti teks yang menggunakan prinsip probabilitas untuk tugas klasifikasi. *Naïve Bayes* adalah bagian dari keluarga algoritma pembelajaran generatif (*family of generative learning algorithms*) yang artinya algoritma yang berupaya memodelkan distribusi masukan dari kelas atau kategori tertentu. Tidak seperti pengklasifikasi diskriminatif, seperti regresi logistic, pengklasifikasi yang tidak mempelajari fitur mana yang paling penting untuk membedakan antar kelas.

Dalam statistik, pengklasifikasi *Naïve Bayes* adalah keluarga “pengklasifikasi probabilistik” linier yang mengasumsikan bahwa fitur-fiturnya independen secara kondisional, mengingat kelas targetnya. Kekuatan (*naivety*) asumsi yang memberi nama pada pengklasifikasi. Pengklasifikasi ini adalah salah satu model jaringan Bayesian yang paling sederhana.

Classifier Naïve Bayes sangat skalabel, memerlukan sejumlah parameter linier dalam jumlah variable (fitur/prediktor) dalam

suatu masalah pembelajaran. Pelatihan kemungkinan maksimum (*maximum likelihood*) dapat dilakukan dengan mengevaluasi ekspresi bentuk tertutup (*closed-form expression*) yang memerlukan waktu linier, bukan pendekatan berulang (*iterative approximation*) yang mahal seperti yang digunakan untuk banyak jenis pengklasifikasi lainnya.



Gambar 3.1 Contoh Pengklasifikasi Naïve Bayes sebagai Jaringan Bayesian

Dalam literatur statistik, model Naïve Bayes dikenal dengan berbagai nama, termasuk *Simple Bayes* dan *Independence Bayes*. Semua nama tersebut merujuk pada penggunaan teorema Bayes dalam aturan keputusan pengklasifikasi, tetapi *Naïve Bayes* bukanlah metode Bayesian.

Pengantar *Naïve Bayes*

Naïve Bayes adalah teknik sederhana untuk membuat pengklasifikasi: model menetapkan label kelas pada contoh masalah, direpresentasikan sebagai vektor nilai fitur, di mana label kelas diambil dari beberapa himpunan berhingga. Tidak ada algoritma tunggal untuk melatih pengklasifikasi seperti itu, namun ada sekumpulan algoritma yang didasarkan pada prinsip umum: semua pengklasifikasi *Naïve Bayes*



IV

TERM FREQUENCY-INVERS DOCUMENT FREQUENCY (TF-IDF)

Dalam pencarian informasi (*information retrieval*), *tf-idf* kependekan dari *term frequency-inverse document frequency*, adalah ukuran pentingnya sebuah kata bagi sebuah dokumen dalam sebuah koleksi atau *korpus*, disesuaikan dengan fakta bahwa beberapa kata lebih sering muncul secara umum. Seperti *model bag-of-words*, model yang memodelkan sebuah dokumen sebagai sekumpulan kata, tanpa urutan kata. Model ini merupakan penyempurnaan dari *model bag-of-words* yang sederhana, dengan membiarkan bobot kata bergantung pada sisa *korpus*.

Tf-idf sering digunakan sebagai faktor pembobotan dalam pencarian informasi (*information retrieval*), *text mining*, dan pemodelan pengguna (*user modelling*). Sebuah survei yang dilakukan pada tahun 2015 menunjukkan bahwa 83% sistem rekomendasi berbasis teks di perpustakaan digital menggunakan *tf-idf*. Variasi

skema pembobotan *tf-idf* sering digunakan oleh mesin pencari (*search engines*) sebagai alat utama dalam penilaian dan pemeringkatan relevansi dokumen berdasarkan permintaan pengguna (*user query*).

Salah satu fungsi pemeringkatan (*ranking functions*) yang paling sederhana dihitung dengan menjumlahkan *tf-idf* untuk setiap kueri; banyak pemeringkatan yang lebih canggih merupakan varian dari model sederhana ini.

Pengertian TF-IDF

Tf-idf merupakan hasil perkalian dua statistik, yaitu *term frequency* dan *inverse document frequency*. Ada berbagai cara untuk menentukan nilai pasti dari kedua statistik tersebut.

Rumus yang bertujuan untuk menentukan pentingnya kata kunci atau frasa dalam dokumen atau halaman web.

Frekuensi Term (*Term Frequency*)

Term frekuensi, $tf(t, d)$, adalah *term* frekuensi relative t dalam dokumen d ,

$$tf(t, d) = \frac{f_{t,d}}{\sum_{t' \in d} f_{t',d}}$$

Di mana $f_{t,d}$ adalah jumlah mentah dari sebuah *term* dalam sebuah dokumen, yaitu jumlah kali *term* t muncul dalam dokumen d . Bila diperhatikan bahwa penyebutnya hanyalah jumlah total *term* dalam dokumen d (menghitung setiap kemunculan *term* yang sama secara terpisah). Ada berbagai cara lain untuk mendefinisikan frekuensi *term*:

- Hitungan mentah itu sendiri: $tf(t, d) = f_{t,d}$
- Frekuensi Boolean: $tf(t, d) = 1$, jika t terjadi di d dan 0 jika tidak.
- Frekuensi skala logaritmik: $tf(t, d) = \log(1 + f_{t,d})$;
- Frekuensi tambahan, untuk mencegah bias terhadap dokumen yang lebih Panjang, misalnya frekuensi mentah dibagi dengan



V

WEB SCRAPING

Pengertian

Web Scraping, *web harvesting*, atau *web data extraction* adalah *data scraping* yang digunakan untuk mengekstraksi data dari situs web. Perangkat lunak *web scraping* dapat langsung mengakses *World Wide Web* menggunakan *Hypertext Transfer Protocol* atau peramban web. Sementara *web scraping* dapat dilakukan secara manual oleh pengguna perangkat lunak, istilah ini biasanya merujuk pada proses otomatis yang diterapkan menggunakan *bot* atau *web crawler*. Perangkat lunak ini adalah bentuk penyalinan di mana data tertentu dikumpulkan dan disalin dari web, biasanya ke dalam basis data lokal atau pusat, *spreadsheet*, *retrieval* atau analisis nanti.

Scraping pada sebuah halaman *web* melibatkan pengambilan (*fetching*) dan ekstraksi dari halaman tersebut. Pengambilan (*fetching*) adalah mengunduh halaman (yang dilakukan peramban (*browser*) saat pengguna melihat halaman). Oleh karena itu, *web crawling* merupakan komponen utama *web scraping*, untuk mengambil halaman untuk kemudian diproses. Setelah diambil, ekstraksi

dapat dilakukan. Konten halaman dapat diparsing, dicari, dan diformat ulang, dan datanya disalin dalam *spreadsheet* atau dimuat ke dalam basis data. *Web scraping* biasanya mengambil sesuatu dari halaman, untuk menggunakannya dengan tujuan lain di tempat lain. Contohnya adalah menemukan dan menyalin nama dan nomor telepon, perusahaan dan URL-nya, atau alamat *email* ke dalam daftar (*contact scraping*)

Selain *contact scraping*, *web scraping* digunakan sebagai komponen aplikasi yang digunakan untuk pengindeksan web (*web index*), *web mining* dan *data mining*, pemantauan perubahan harga daring dan perbandingan harga (*price comparison*), *scraping* ulasan produk (untuk mengamati persaingan), pengumpulan daftar real estate, pemantauan data cuaca, deteksi perubahan situs web (*website change detection*), riset (*research*), pelacakan kehadiran (*tracking online presence*), reputasi daring (*online reputation*), penggabungan web (*web mashup*) dan integrasi data web (*web data integration*).

Halaman web (*web page*) dibuat menggunakan bahasa *markup* berbasis teks (*text-based mark-up languages*) atau (HTML dan XHTML), dan sering kali berisi banyak data yang berguna dalam bentuk teks. Akan tetapi, Sebagian besar halaman web dirancang untuk pengguna akhir (*end-user*) dan bukan untuk kemudahan penggunaan otomatis. Akibatnya, alat dan perangkat lunak khusus telah dikembangkan untuk memfasilitasi *scraping* halaman web (*web of page scraping*). Aplikasi *web scraping* mencakup riset pasar (*market research*), perbandingan harga (*price comparison*), pemantauan konten (*monitoring content*), dan masih banyak lagi. Bisnis mengandalkan layanan *web scraping* untuk mengumpulkan dan memanfaatkan data secara efisien.

Bentuk *web scraping* yang lebih baik melibatkan pemantauan umpan data server web. Misalnya JSON umumnya digunakan sebagai mekanisme transportasi antara klien dan server web.

Ada beberapa metode yang digunakan beberapa situs web untuk mencegah *web scraping*, seperti mendeteksi dan melarang *bot crawling*



VI

MEDIA SOSIAL TWITTER/X

*T*witter adalah media sosial daring (*social networking service*) dan layanan jejaring sosial (*social media website*). Twitter adalah salah satu situs web media (*social media website*) terbesar di dunia dan salah satu web yang paling banyak dikunjungi di dunia. yang dioperasikan oleh perusahaan Amerika Serikat, *Twitter, Inc.* Perusahaan yang berganti nama menjadi *X* mulai Juli 2023 dan dengan nama *X Corp.* Pengguna dapat berbagi pesan teks pendek, gambar, dan video dalam kiriman (*posts*) yang sebelumnya dikenal dengan “*tweet*” dan *like* (menyukai) atau *memposting* ulang atau *me-retweet* konten pengguna lain. *Twitter* juga mencakup pesan langsung (*direct messaging*), panggilan video dan audio, *bookmark*, daftar dan komunitas, *Spaces*, fitur audio sosial. Pengguna dapat memberikan suara pada konteks yang ditambahkan oleh pengguna yang disetujui menggunakan fitur Catatan Komunitas (*Community Notes*).

Twitter diciptakan pada Maret 2006 oleh Jack Dorsey, Noah Glass, Biz Stone, dan Evan Williams, dan diluncurkan pada Juli 2006. *Twitter, Inc.*, berkantor pusat di San Francisco, California dan memiliki lebih dari 25 kantor di seluruh dunia. Ciri khas layanan ini

pada awalnya adalah bahwa kiriman (*posts*) harus singkat. Kiriman (*posts*) awalnya dibatasi hingga 140 karakter, yang diubah menjadi 280 karakter pada tahun 2017. Batasan tersebut dihapus untuk akun yang berlangganan pada tahun 2023. Mayoritas *tweet* dihasilkan oleh minoritas pengguna.

Penampilan dan Fitur (*Appearance and Features*)

1. Tweet

Tweet dapat dilihat publik secara *default*, tetapi pengirim dapat membatasi pengiriman pesan hanya kepada pengikutnya saja. Pengguna dapat menonaktifkan pengguna yang tidak ingin diajak berinteraksi, memblokir akun agar tidak melihat kirimannya, dan menghapus akun dari daftar pengikutnya.

Pengguna dapat mengirim melalui situs web *twitter*, aplikasi eksternal yang kompatibel (seperti *smart phone*), atau melalui layanan pesan singkat (SMS) yang tersedia di negara-negara tertentu. Pengguna dapat menyetujui (*subscribe*) kiriman pengguna lain, yang dikenal sebagai “mengikuti” (“*following*”) dan pengguna menyetujui (*subscribers*) dikenal sebagai “pengikut” (“*followers*”) atau “*tweeps*”, gabungan dari *twitter* dan *peeps*. Kiriman individual dapat diteruskan oleh pengguna lain ke umpan baliknya sendiri, proses yang disebut “*repost*” atau “*retweet*”

2. Konten Multimedia

Pada tahun 2016, *twitter* mulai lebih besar berfokus pada pemrograman video streaming langsung, menyelenggarakan berbagai acara. *Twitter* mengumumkan sejumlah kemitraan (*partnerships*) baru dan yang diperluas untuk layanan video streaming pada acara, antara lain *Bloomberg*, *BuzzFeed*, *Cheddar*, *IMG Fashion*, *Live Nation Entertainment*, *Major League Baseball*, *MTV* dan *BET*, *NFL Network*, *PGA Tour*, *The Players’ Tribune*, *Ben*



VII

WORD CLOUD DAN UNDANG-UNDANG CIPTA KERJA

Pengertian

Word cloud merupakan bentuk representasi visual dari kumpulan kata-kata di mana kata-kata yang paling sering muncul dalam teks ditampilkan dalam ukuran yang lebih besar dan menonjol, sedangkan kata-kata yang kurang umum muncul dalam ukuran yang lebih kecil. *Word cloud* digunakan untuk memvisualisasikan frekuensi kemunculan kata-kata dalam sebuah teks atau dokumen dengan cara yang lebih menarik dan mudah dimengerti.

Word cloud memiliki beberapa kegunaan yang bermanfaat dalam berbagai konteks, diantaranya:

1. Visualisasi data teks

Word cloud dapat digunakan untuk mengidentifikasi kata-kata kunci atau istilah yang paling sering muncul dalam teks.



VIII

MENELISIK SENTIMEN UU CIPTA KERJA DI TWITTER

Penyelidikan sentimen undang-undang cipta kerja adalah suatu cara ilmiah untuk mendapatkan data yang valid dengan tujuan dapat ditemukan, dikembangkan, atau dibuktikan, suatu pengetahuan tertentu sehingga pada gilirannya dapat digunakan untuk memahami, memecahkan, dan mengantisipasi masalah dalam bidang tertentu.

Pengumpulan Data

Teknik pengumpulan data menggunakan *web scraping* yaitu teknik mengambil data dalam jumlah besar dari berbagai website secara otomatis. Cara ini membantu mengekstrak data yang relevan dan bermanfaat, yang kemudian dapat digunakan untuk berbagai keperluan, seperti analisis data maupun *market research* secara spesifik.

Data yang digunakan adalah data *tweet* pada *twitter* tentang Undang-Undang Cipta Kerja merupakan *Omnibus Law* dalam hukum

Indonesia yang disahkan oleh Pemerintah. Teknik pengumpulan data menggunakan teknik *scraping* data *tweet* pada *twitter* menggunakan *library Twint*. Data yang dikumpulkan merupakan data *tweet* yang mengandung kata dan *hashtag* “Undang-Undang Cipta Kerja” dan “*Omnibus Law*”. Data yang diambil adalah data sejumlah 1000 *tweet* yang dikirim pada bulan Oktober 2020 hingga November 2020.

Pelabelan Manual

Teknik *scraping* web manual adalah *copy and paste human* merupakan bentuk web *scraping* yang paling sederhana adalah menyalin (*copy*) dan menempel (*paste*) data secara manual dari halaman web ke dalam file teks atau *spreadsheet*. Terkadang, bahkan teknologi web *scraping* terbaik tidak dapat menggantikan pemeriksaan manual dan *copy-paste* manusia, dan mungkin satu-satunya solusi yang dapat diterapkan ketika situs web untuk *scraping* secara eksplisit menyiapkan penghalang untuk mencegah otomatisasi mesin.

Data yang telah dikumpulkan selanjutnya disaring (diseleksi) untuk mendapatkan data sesuai berdasarkan kebutuhan. Proses penyaringan (seleksi) dilakukan dengan menyaring data yang tidak mengandung opini. Data diperoleh dari hasil seleksi adalah sebanyak 687 data.

Pemberian label pada data dilakukan secara manual ke dalam klasifikasi kelas yang telah ditentukan, yaitu negatif dan positif. Hasil dari pelabelan secara manual diperoleh 357 *tweet* dengan kelas positif dan 330 *tweet* dengan kelas negatif.

Preprocessing

Data *preprocessing* adalah proses mengubah data mentah ke dalam bentuk yang lebih mudah dipahami. Proses ini diperlukan untuk memperbaiki kesalahan pada data mentah yang seringkali dianggap tidak lengkap dan memiliki format yang tidak teratur.

Preprocessing melibatkan proses validasi dan imputasi data. Validasi bertujuan untuk menilai tingkat kelengkapan dan akurasi



IX

HASIL ANALISIS SENTIMEN UU CIPTA KERJA

Penghimpunan Data

Penghimpunan data *tweet* dilakukan dengan menggunakan teknik *web scraping* dengan memanfaatkan *library Twint*. Data yang dikumpulkan merupakan data *tweet* yang mengandung kata “Undang-Undang Cipta Kerja” dan *hashtag* “*Omnibus Law*”. Data *tweet* yang dikumpulkan sejumlah 1000 *tweet*. Berikut ini merupakan sampel *tweet* yang diperoleh

	id	conversation_id	created_at	date	timezone	place	tweet	language	hashtags	cashtags	...	geo	source	user_rt_id
0	1324137994312421377	1324137994312421377	1.604534e+12	2020-11-05 06:54:33	700	NaN	Cavalkot Alkhyar Tolak Omnibus Law https://t.c...	in			NaN	NaN	NaN	NaN
1	1324137112342441985	1324137112342441985	1.604534e+12	2020-11-05 06:51:03	700	NaN	Siapkan dirimu dan rapatkan barisan perlawanan...	in	['tolakomnibuslaw', 'panjangumpurjuangan']		NaN	NaN	NaN	NaN
2	1324136805252231168	1324136805252231168	1.604534e+12	2020-11-05 06:49:50	700	NaN	Yg paling bertanggungjawab itu pasti bos bebek...	in	['omnibuslaw']		NaN	NaN	NaN	NaN

Gambar 9.1. Sampel *tweet* hasil *web scraping*

Pelabelan Manual

Data yang dikumpulkan dengan teknik *web scraping* kemudian dirapikan yaitu dengan menghapus kolom-kolom yang tidak diperlukan. Data yang telah dikumpulkan selanjutnya disaring (diseleksi). Data disaring dengan menghapus *tweet* yang tidak berisi opini. Data diperoleh dari hasil seleksi adalah sebanyak 687 data.

Adapun sampel data *tweet* yang telah diberi label seperti yang ditampilkan pada tabel 9.1 berikut:

Tabel 9.1. Data yang telah diberi label

No	<i>Tweet</i>	Sentimen
1	Pemerintah membentuk pengelola investasi guna kemajuan UMKM dan peningkatan jumlah lapangan kerja #UUCiptaKerja#OmnibusLaw https://t.co/TIk5mK5bwS	Positif
2	@Yusrilihza_Mhd Sepertinya UU Celaka/omnibus law merupakan UU yang di sahkan& di tangani Secara cacat Hukum, karena; pengesahan & penandatanganan nya pada UU yg belum siap/belum selesai dan juga UU tersebut tidak di terima lebih dari separuh lebih dari pada Bangsa Indonesia	Negatif
3	UU ini bertujuan mewujudkan masy.Indonesia Makmur, adil dan sejahtera. Poin mana saja yg kalian tolak ? Sementara disisi lain sudah berkontribusi sampai dimana kalian untuk mewujudkan masyarakat yg kita dambakan. Tetap optimis dan lihat UU Omnibus Law Cipta Kerja ini bekerja	Positif
4	@republikaonline. @OssyDermawan Kenapa Presiden tanda tangan juga sdh tau temukan pasal 6 UU Cipta Kerja Bermasalah, Aneh begini yah apakah sdh ada titipan dari investor ! sejak ketok palu Azis yg pimpin rapat UU Cipta Kerja dan mattiin Mic oleh Puan tdk memberi peluang masukan dari peserta rapat ! aneh.....	Negatif
5	UU Ciptakerja sudah ditandatangani, Jrx divonis penjara 3 tahun, padahal dia dengan mempertaruhkan 'panggung besar' nya, berani bersuara banyak tentang isu lingkungan yang dianggap 'sepele':#Balitolakreklame #Jogjaoradidol #savepetanikendeng Yakini negeri ini baik baik saja??	Negatif



X

PENUTUP

Implementasi metode *Naïve Bayes Classifier* dan *Term Frequency-Inverse Document Frequency* (TF-IDF) yang digunakan untuk melakukan analisis sentimen tentang Undang-Undang Cipta Kerja pada media sosial *Twitter* dengan menggunakan data *tweet* yang mana data yang digunakan sebagai *data training* sebesar 80% dan data testing sebesar 20% menghasilkan nilai *accuracy* sebesar 89.9%, *precision* sebesar 90%, *recall* sebesar 89.9% dan *f-1 score* sebesar 89.9%. Analisis sentimen masyarakat Indonesia terhadap Undang-Undang Cipta Kerja yang divisualisasikan melalui media sosial *Twitter* menggunakan *data testing* sebesar 138 data, berdasarkan hasil klasifikasi menunjukkan bahwa masyarakat Indonesia 52.9% menyatakan kontra dan 47.1% menyatakan pro terhadap Undang-Undang Cipta Kerja;

Hasil analisis *word cloud* sentimen negatif dan sentimen positif menunjukkan bahwa terdapat persamaan dan perbedaan pada sentimen. Persamaannya adalah banyaknya kata Jokowi, pemerintah, pekerja, buruh, yang paling sering pada *tweet* pada kedua sentimen tersebut. Hal ini menunjukkan pihak yang paling sering muncul pada

tweet tentang Undang-Undang Cipta Kerja. Sedangkan perbedaan dari kedua sentimen tersebut adalah kata-kata yang memiliki makna yang berlawanan dan bertolak belakang. Dari persamaan dan perbedaan sentimen tersebut menunjukkan persepsi pengguna *Twitter* terhadap Undang-Undang Cipta Kerja



INDEKS

A

- Accuracy 79, 95
- Agregasi Vertikal 40
- AI-Powered Document Understanding 41
- Akun Bot 60
- Akun mencurigakan dan disengketakan 60
- Akun Pengguna 59
- Akun Terverifikasi 59
- Alat Web Scraping yang Populer 42
- Algoritma (Algorithm) 10, 11, 23, 40, 50, 100, 102, 111, 112
- Analisis Halaman Web Visi Komputer 41
- Analisis sentimen 1, 2, 7, 8, 9, 12, 93, 95
- Antarmuka 49, 56, 58
- API Dan Platform Pengembang 57
- Apify 43, 95
- Appearance and Features 48
- Aspect-Based 12

B

- Bayesian Probability 17
- BeautifulSoup 42, 95
- Bernoulli 19, 20, 22, 96
- Boolean 22, 30, 95
- Bot 43, 44, 60, 95
- Bot Accounts 60
- business process automation (BPA) 73, 95
- Buy Button 52

C

- CAPTCHA 44
- case folding 73, 80, 85, 96
- Catatan Komunitas 47, 62
- Censorship and Moderation 62
- Chatbot 96
- Cleaning 73, 85, 96
- Computer Vision Web-Page Analysis 41
- confusion matrix 78, 79, 88
- Content 54
- Copy and Paste 39
- Crawling 96
- CSS 44, 98

D

Dampak 61
data testing 77, 78, 87, 88, 89, 93, 102
data training 76, 77, 78, 87, 88, 93, 102
deep learning 10
Demografi 53, 55
Derivatif 36
Detailed Assesment 12
Document Classification 25
DOM parsing 39, 97

E

E-commerce 52
Emotion Detection 12
Estimasi Parameter 19, 23
Evaluasi Performa 78, 88
event model 19, 22

F

Fetching 97
Filtering 74, 85, 97
Fleksibilitas dalam Desain 65

G

Gangguan 58
Gaussian 20, 98

H

Harassment 59, 98
HTTP Programming 39
Hybrid 11

Hybrid Approach 11

I

Identifikasi Kata Kunci 65
Impact 61
Implementasi 56, 93, 112
information retrieval 29, 99
Innovators Paten Agreement 57
InstantAPLai 43, 99
Interface 42, 58, 95, 113
inverse document frequency 29, 30, 31

K

Keamanan 58, 59, 60
Kekurangan Informasi Lengkap 66
Kelemahan word cloud 65
Kesulitan dalam Penanganan Data Besar 66
Keterlibatan Pengguna 55
Klasifikasi Dokumen 25
Konten 38, 48, 54, 55, 61, 65
Konten Multimedia 48
Kurasi 49

L

Levels of Use 55
library Twint 72, 83, 99
linier regression 10
Logistic Regression 23

M

Machine Learning 9, 10, 100, 112

Mallcious and Fake Accounts 60

Masyarakat 61

Matriks evaluasi 79

Maximum a Posteriori (MAP) 19, 100

maximum likelihood 16, 17, 99

media sosial iii, iv, 2, 4, 7, 8, 9, 13, 47, 64, 93, 98

Mengidentifikasi Tren dan Pola 64

Meningkatkan Pemahaman Konten 65

Merangkul Informasi 64

Metrik evaluasi 77

Mobile 51

Mode Keamanan 59, 60

model bag-of-words 29, 100

Model Kejadian 19

Model Probabilistik 17

Model Probabilitas 19

Moderasi Tweet 62

Momen (Moments) 50

Monetisasi Pengguna 52

N

Naïve Bayes 2, 10, 15, 16, 17, 19, 20, 21, 22, 23, 24, 25, 27, 28, 76, 78, 80, 87, 88, 89, 93, 98, 100, 112

Natural Language Processing (NLP) 1, 7, 9, 39, 95, 100

Negasi 13, 100

Normalisasi 33, 74, 85, 100

O

Octoparse 43, 101

Omnibus Law 2, 71, 72, 83, 84, 101, 112

Open Source 57

Open-source software 56, 101

Opini 2, 3, 101, 112

Outages 58, 101

Overfitting 78, 101

P

ParseHub 43, 101

Parsing 40, 101

Parsing HTML 40

Pelabelan Manual 72, 84

Pemahaman Dokumen Kekuatan AI 41

Pembagian data 77

Pembenaran idf 33

Pembobotan 32, 33

Pemodelan 76, 77

Pemrograman HTTP 39

Pencocokan Pola Teks 39

Pengenalan Anotasi Semantik 41

Penggunaan 10, 13, 53, 55, 61, 67

Pengujian Pemodelan 77

- Pengumpulan Data** 71
- Penguraian Document Object Model** 40
- Penguraian HTML** 40
- Perangkat Lunak** 42
- Perdagangan Elektronik** 52
- Perjanjian Paten Inovator** 57
- Platform** 40, 43, 57, 101
- Platform Web Scraping** 43
- Popular Web Scraping Tools** 42
- Precession** 79
- Preprocessing** 9, 72, 85
- Pribadi (Privacy)** 59
- probabilitas bersyarat** 17, 18, 28, 96
-
- R**
- Recall** 79, 102
- Regresi Logistik** 23
- Retweet** 51, 102
- Ringkasan Informasi** 65
-
- S**
- scraping** 37, 38, 39, 41, 42, 43, 44, 45, 71, 72, 83, 84, 95, 96, 98, 101, 104, 113
- Scrapy** 43, 102
- Security** 58
- Selular** 51
- Semantic Annotation Recognizing** 41
- semi-supervised** 23
- Semi-Supervisi** 23
- sentimen positif** 1, 2, 4, 11, 87, 89, 91, 92, 93
- sentiment analysis** iii, 1, 95
- Shop Module** 52, 102
- Society** 61
- Software** 42, 57
- spam** 25, 26, 27, 54, 60, 102
- split validation** 78, 87
- stemming** 73, 80, 85
- stopword** 66, 75, 85, 103
- streaming** 48, 53
- Subjektivitas Interpretasi** 65
- subscriptions** 52
- Sumber Terbuka** 57
- Super Follows** 52
- supervised machine learning** 15, 76, 102
- support vector machine** 10
- Suspect and Contested Accounts** 60
-
- T**
- Teknik Web Scraping** 39
- Teknologi** 8, 9, 56, 111, 112, 115
- term frequency** 29, 30
- Term Frequency-Inverse Document Frequency** 32, 33, 75, 76, 86, 93, 103
- Term Weighting** 75, 80
- Text Analysis** 7
- Text Pattern Machine** 39
- TF-IDF** 29, 30, 75, 76, 78, 86, 87, 88, 93, 103, 112

Ticketed Spaces 52, 103
Tingkat Penggunaan 55
Tip Jar 52, 103
Token 73
Tokenisasi 9, 74, 103
Tokenization 73, 85
train test split 77, 103
Tweet 48, 56, 59, 62, 84, 85, 104
Twitter iii, iv, 2, 4, 5, 47, 48, 49,
50, 51, 52, 53, 54, 56, 57, 58,
59, 60, 61, 62, 93, 94, 104,
105, 111, 112, 113
Twitter Blue 51
Twitter Spaces 49, 104

U

Underfitting 104
Usage 53, 61
User Accounts 59
User Engagement 55

V

Validasi model 77
Verified Accounts 59
Verifikasi akun berbayar 52
Vertical Aggregation 40, 104
Visualisasi data teks 63
Visualisasi Menarik 65

W

Web Scraping 37, 39, 42, 43, 113
word cloud 64, 65, 66, 67, 90,
91, 92, 93
word cloud yang efektif 66



DAFTAR PUSTAKA

- R Indriati, R Firliana, (2014). “*Opinion Mining Menggunakan Hidden Markov Model dan Algoritma Viterbi*”, *Nusantara of Engineering (NOE)* 1 (1), <https://ojs.unpkediri.ac.id/index.php/noe/article/view/34>
- R Indriati, AR Amna, (2014). “*Mining Skill Dan Keahlian Dari Teks Menggunakan Ekstraksi Informasi*”, *Nusantara of Engineering (NOE)* 1 (2), <https://ojs.unpkediri.ac.id/index.php/noe/article/view/64>
- A Pramono, R Indriati, A Nugroho, (2017) “*Sentimen Analisis Tokoh Politik Pada Twitter*”, *Prosiding SEMNAS INOTEK (Seminar Nasional Inovasi Teknologi)* 1 (1), 195-200, ISBN: 978-602-61393-0-6, e-ISSN: 2549-7952
- JM Suharto, R Indriati, T. Andriyanto (2017). “*Analisis Perilaku Konsumen Pada Pembelian Produk Perlengkapan Bayi*”, *Prosiding SEMNAS INOTEK (Seminar Nasional Inovasi Teknologi)* 1 (1), 295-300, ISBN: 978-602-61393-0-6, e-ISSN: 2549-7952
- R Indriati, T Andriyanto, (2017). “*RESTFul Web Service Untuk Integrasi Antar Sistem Informasi*”, *Jurnal Sains dan Informatika* 3 (2), 116-121, p-ISSN: 2460-173X, e-ISSN: 2598-5841, doi:<https://doi.org/10.34128/jsi.v3i2.111>
- AG Tammam, S Sucipto, R Indriati, (2018). “*Hoax Detection at Social Media With Text Mining Clarification System-Based*”, *JUPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)* 3 (2), 94-100, e-ISSN: 2450-8984, doi:10.29100/jupi.v3i2.837
- IY Anggraini, S. Sucipto, R Indriati, (2018). “*Cyberbullying Detection Modelling at Twitter Social Networking*”, *JUITA:*

Jurnal Informatika 6 (2), 113-118, p-ISSN:2086-9398,
e-ISSN:2579-9801, doi:10.30595/juita.v6i2.3350

- R. Indriati, SB Persada, (2018). “*Information Retrieval* Pencarian Informasi Jenis Musik”, Prosiding SEMNAS INOTEK (Seminar Nasional Inovasi Teknologi), 2 (1), 305-310, e-ISSN: 2549-7952, p-ISSN:2580-3336
- R Indriati, TN Wijaya, MN Muzaki, (2021). “Analisis Sentimen Opini Publik Tentang Undang-Undang Cipta Kerja Pada *Twitter*”, Jambura Journal of Electrical and Electronics Engineering 3 (2), 78-83, e-ISSN: 2715-0887, p-ISSN: 2654-7813, doi: <https://doi.org/10.37905/jjee.v3i2.10885>
- R Indriati, (2023). “Pengantar Sistem Informasi”, Universitas Nusantara PGRI Kediri, 1-184, ISBN: 978-623-8330-12-6
- E.N. Hamdana, (2021). “Pengembangan Sistem Analisis Sentimen Berbasis Java Data *Twitter* Terhadap Omnibus Law Menggunakan Algoritma Naïve Bayes dan K-Nearest Neighbor (K-NN)”, J. Inform. Polinema vol 7 No 2, hal. 79-84, 2021, doi: 10.33795/jip.v7i2.688
- A. Hasan, S. Moin, A. Karim, dan S. Shamshirband, (2018). “*Machine Learning-Based Sentiment Analysis for Twitter Accounts*” Math. Comput. Appl, vol 23 no 1, hal. 11
- M. I. Fikri, T. S. Sabrila, dan Y. Azhar, (2020). “Perbandingan Metode Naïve Bayes dan Support Vector Machine pada Analisis Sentimen *Twitter*”, Smatika J., vol 10 no 02, hal. 71-76
- M. A. Rofiqi, A. C. Fauzan, A. P. Agustin, dan A. A. Saputra, (2019). “Implementasi Term-Frequency Inverse Document Frequency (TF-IDF) Untuk Mencari Relevansi Dokumen Berdasarkan Query”, Ilk. J. Comput. Sci. Appl. Informatics, vol 1 no. 2, hal. 28-64, <https://doi.org/10.28926/ilkomnika.v1i2.18>
- Mayur Wankhade; Annavarapu Chandra Sekhara Rao; Chaitanya Kulkarni; (2022). “A Survey On Sentiment Analysis Method, Applications, and Challenges”, Springer, Artificial Intelligence Review (2022) 55:5731-5780, <https://doi.org/10.1007/s10462-022-10144-1>
- Thapelo, Tsaone Swaabow; Namoshe, Molaletsa; Matsebe, Oduetse; Motshegwa, Tshiamo; Bopape, Mary-Jane Morongwa (2021-

07-28). "SASSCAL WebSAPI: A Web Scraping Application Programming Interface to Support Access to SASSCAL's Weather Data". *Data Science Journal*. 20:24. Doi:10.5334/dsj-2021-024. ISSN 1683-1470.S2CID 237719804

https://en.wikipedia.org/wiki/Machine_learning, 4 Desember 2023

<https://aws.amazon.com/what-is/sentiment-analysis>, diakses 30 Januari 2024

<https://brand24.com/sentiment-analysis-and-listening>, diakses 3 Februari 2024

<https://monkeylearn.com/sentiment-analysis>, diakses 4 Februari 2024

<https://www.ibm.com/id-id/topics/sentiment-analysis>, diakses 20 Maret 2024

<https://www.ibm.com/topics/naive-bayes>, diakses 24 Juli 2024

https://en.wikipedia.org/wiki/Naive_Bayes_classifier, diakses 29 Juli 2024

<https://www.ibm.com/topics/naive-bayes>, diakses 2 Agustus 2024

<https://www.geeksforgeeks.org/naive-bayes-classifiers/>, 1 Agustus 2024

<https://en.wikipedia.org/wiki/Tf>, diakses 18 Juli 2024

[https://id.wikipedia.org/wiki/X_\(media_sosial\)](https://id.wikipedia.org/wiki/X_(media_sosial)), diakses 25 Agustus 2024

<https://en.wikipedia.org/wiki/Twitter>, diakses 13 Agustus 2024

<https://peraturan.bpk.go.id/Details/149750/uu-no-11-tahun-2020> diakses 24 Juli 2024

RUU Cipta Kerja.PDF, download 9 Agustus 2024

<https://peraturan.bpk.go.id/Details/149750/uu-no-11-tahun-2020> diakses 9 Agustus 2024

<https://revou.co/kosakata/web-scraping>, diakses 12 Juni 2024

<https://dcloudhost.com/blog/mengenal-apa-itu-data-atau-web-scraping>, diakses 30 Juli 2024

<https://revou.co/kosakata/web-scraping#:>, diakses 5 Agustus 2024

https://en.wikipedia.org/wiki/Web_scraping, diakses 15 Agustus 2024

PROFIL PENULIS



Tentang Saya:

- Dosen Fakultas Teknik dan Ilmu Komputer Program Studi Sistem Informasi Universitas Nusantara PGRI Kediri
- Profesional Teknologi Informasi fokus: Pendidikan Sistem Informasi Manajemen, Konsultan, Elektronik Pemerintahan, basisdata, Sentimen Analisis, Kecerdasan Buatan (AI).
- Entrepreneurship

Rini Indriati, M. Kom; dosen Fakultas Teknik dan Ilmu Komputer Program Studi Sistem Informasi Universitas Nusantara PGRI Kediri. Menyelesaikan studi program Sarjana Komputer dari STIMIK STIKOM Surabaya, Pasca Sarjana Jurusan Teknologi Informasi Sekolah Tinggi Teknik Surabaya.

Pada Program Studi Sistem Informasi aktivitas sebagai dosen yang mengampu mata kuliah Pengantar Sistem Informasi, Sistem Informasi Manajemen, Analisa Proses Bisnis, Manajemen Layanan teknologi Informasi

Dosen yang telah menerbitkan buku referensi, buku ajar dan beberapa jurnal yang relevan dengan keilmuan yang diperoleh pada bidang ilmu teknologi informasi.

Buku ini merupakan hasil kajian yang berfokus pada bidang ilmu Analisis Sentimen (sentiment analysis), suatu metode yang semakin berkembang pesat dalam era digital. Analisis Sentimen mengacu pada proses analisis teks digital yang bertujuan untuk menentukan apakah suatu teks memiliki muatan emosional positif atau negatif. Metode ini sangat berguna dalam memahami pandangan dan persepsi publik terhadap berbagai topik, salah satunya adalah kebijakan pemerintah.

Dalam monograf ini, kami menggunakan platform media sosial Twitter sebagai sumber pengambilan teks digital. Twitter dipilih karena popularitasnya yang tinggi di kalangan masyarakat Indonesia. Sebagai salah satu aplikasi media sosial terbesar, Twitter memungkinkan pengguna untuk mengirimkan dan menerima pesan secara real-time, yang menjadikannya sebagai ruang publik yang dinamis dan efektif dalam mengukur opini serta persepsi masyarakat. Melalui karakteristik ini, Twitter menjadi media ideal untuk memonitor sentimen publik terhadap kebijakan besar seperti Undang-Undang Cipta Kerja.

Proses kajian dalam buku ini mencakup pemanfaatan teknik analisis sentimen untuk menelaah bagaimana masyarakat bereaksi terhadap UU Cipta Kerja melalui platform Twitter. UU Cipta Kerja merupakan kebijakan yang memunculkan diskusi luas dan beragam pandangan di berbagai kalangan.

Analisis Sentimen
**UNDANG-UNDANG
CIPTA KERJA**
PADA TWITTER

