



5.19%

SIMILARITY OVERALL

SCANNED ON: 23 JUL 2024, 12:31 PM

Similarity report

Your text is highlighted according to the matched content in the results above.

IDENTICAL 0.35% **CHANGED TEXT** 4.84% **QUOTES** 29.3%

Report #22128215

BAB I PENDAHULUAN 1.1 Latar Belakang Masalah Kendaraan listrik semakin mendapatkan perhatian yang meningkat di seluruh dunia sebagai alternatif ramah lingkungan dalam rangka mengatasi masalah perubahan iklim dan ketergantungan terhadap bahan bakar fosil. **10** Di Indonesia, pemerintah juga telah memperkenalkan kebijakan subsidi kendaraan listrik sebagai bagian dari upaya untuk mengurangi emisi gas rumah kaca dan meningkatkan efisiensi energi.

Subsidi kendaraan listrik diharapkan dapat mendorong adopsi teknologi ini dan berkontribusi terhadap pembangunan berkelanjutan. Namun, dalam implementasi kebijakan subsidi kendaraan listrik, terdapat dinamika kompleks yang melibatkan berbagai aspek termasuk ekonomi, lingkungan, dan sosial. Salah satu aspek yang tidak kalah penting adalah sikap serta respons masyarakat terhadap kebijakan ini. Persepsi dan sentimen masyarakat terhadap kebijakan subsidi kendaraan listrik dapat memiliki dampak signifikan terhadap efektivitas dan penerimaan kebijakan ini dalam masyarakat. Dalam menghadapi tantangan ini, analisis sentimen masyarakat dapat menjadi alat penting untuk memahami opini dan pandangan yang beredar di kalangan masyarakat terkait dengan kebijakan tersebut. Dengan menganalisis sentimen, kita dapat mengidentifikasi apakah masyarakat cenderung mendukung, atau menentang kebijakan subsidi kendaraan listrik. Penggunaan algoritma seperti Long Short-Term Memory (LSTM) dalam analisis sentimen dapat membantu dalam memahami kompleksitas analisis teks yang

berhubungan dengan kebijakan. **9** Oleh karena itu, penelitian ini bertujuan untuk mengembangkan aplikasi analisis sentimen yang dapat menganalisis respon dan sentimen masyarakat terhadap kebijakan subsidi kendaraan listrik di Indonesia.

Pengembangan aplikasi ini diharapkan dapat memberikan wawasan berharga bagi para pengambil keputusan, pemerintah, dan pihak-pihak terkait dalam mengukur efektivitas kebijakan serta merespon kebutuhan dan harapan masyarakat terkait kendaraan listrik. 1.2 Identifikasi Masalah Peneliti merumuskan masalah dan batasan masalah dari sudut pandang peneliti.

Rumusan masalah dapat digunakan menjadi acuan mengenai masalah yang diangkat 1.2 **15** **1**

Rumusan Masalah Berdasarkan uraian masalah pada latar belakang, peneliti merumuskan masalah yaitu sebagai berikut. 1. Bagaimana membangun model analisis sentimen terhadap sistem teknologi dan kebijakan subsidi kendaraan listrik di media sosial Twitter menggunakan pendekatan pengolahan bahasa alami? 2. Bagaimana membangun aplikasi untuk melihat hasil analisis sentimen berdasarkan input yang diberikan oleh user. 3.

Bagaimana perbandingan sentimen mengenai teknologi dan kebijakan subsidi kendaraan listrik di media sosial Twitter pada awal perilisian desember

2022 sampai agustus 2023 1.2 **11** **2** Batasan Masalah Batasan penelitian

dilakukan untuk mempersempit masalah dan tidak keluar pembahasan yang telah dijabarkan pada rumusan masalah. Adapun Batasan-batasan pada penelitian

ini yaitu sebagai berikut. a. Data yang digunakan yaitu tweet yang mengandung kata kunci subsidi kendaraan listrik dengan filter Bahasa Indonesia. b. Format data pada unggahan file yaitu comma separated

value atau csv. c. Kategori sentimen yang digunakan pada penelitian ini yaitu negatif dan positif. 1.3 Tujuan Penelitian Tujuan dari

penelitian ini yaitu mengetahui sentiment masyarakat tentang teknologi dan kebijakan subsidi kendaraan listrik yang bersumber dari media

sosial Twitter, melihat perbedaan sentimen publik terhadap kebijakan subsidi kendaraan listrik dari data saat awal rilis pada tahun desember 2022 dan hingga agustus 2023. Hasil dari model klasifikasi

ini bertujuan untuk menyimpulkan informasi dari data- data yang



tersebar di media sosial Twitter dan menyajikan hasilnya sebagai masukan bagi pemerintah atau pihak yang berwenang untuk melakukan evaluasi terhadap program yang saat ini berjalan. 1.4 Manfaat Penelitian Penelitian ini diharapkan mampu memberikan manfaat kepada Peneliti, Ilmu pengetahuan, dan Masyarakat. 1.4.1 Manfaat untuk Masyarakat/calon pengguna Penelitian ini diharapkan dapat bermanfaat bagi masyarakat yang ingin mengetahui opini publik mengenai kebijakan subsidi kendaraan listrik. Aplikasi analisis sentimen diharapkan dapat membantu masyarakat yang ingin mengetahui sentimen di Twitter menggunakan dataset yang diunggah oleh pengguna. 1.4.2 Manfaat Untuk Ilmu Pengetahuan Peneliti berharap bahwa penelitian ini mampu memberi informasi tentang analisis sentimen terhadap kebijakan subsidi kendaraan listrik dengan pendekatan pengolahan bahasa alami, serta menambah referensi terkait analisis sentiment terhadap kebijakan subsidi kendaraan listrik. 1.5 Kebaruan Kebaruan penelitian ini terletak pada kasus yang diangkat yaitu mengetahui sentimen masyarakat Indonesia terhadap kebijakan subsidi kendaraan listrik pada media sosial Twitter. Model NLP akan dikemas dalam aplikasi dan dapat digunakan untuk melakukan analisis sentimen sesuai masukan yang diberikan user. 12 1.6 Sistematika Penulisan Sistematika penulisan merupakan subbab untuk memahami sistematika penelitian dalam laporan ini, uraian babnya adalah sebagai berikut : 1. 6 Bab I Pendahuluan Bab ini menjelaskan latar belakang masalah, identifikasi masalah, batasan masalah, tujuan penelitian, manfaat penelitian, kebaruan, kerangka penelitian, dan target luaran dari masalah yang diangkat. 2. Bab II Tinjauan Pustaka Bab ini berisi pencapaian terdahulu dan tinjauan teoritis yang memuat berbagai teori yang menjadi dasar pengetahuan dalam melakukan penelitian dan penulisan laporan akhir. Bab ini membahas tentang hasil-hasil penelitian yang terdahulu, konsep dan teori yang berhubungan dengan masalah yang dianalisis dalam penelitian, serta kajian-kajian yang terkait dengan permasalahan yang akan diteliti. Tujuan dari bab ini adalah memberikan gambaran tentang konsep dan teori yang berhubungan

dengan masalah, serta memperoleh wawasan tentang kajian-kajian yang terdahulu dan memastikan penelitian ini memiliki kebaruan. 3. Bab III Tahapan Pelaksanaan Penelitian Bab ini membahas metode yang digunakan dalam penelitian untuk merealisasikan sistem. Bab ini terdiri dari langkah-langkah pelaksanaan, metode pengembangan sistem, dan metode pengujian. 4. Bab IV Perancangan Pada bab ini menjelaskan tentang analisis sistem terdahulu, spesifikasi kebutuhan sistem baru, perancangan sistem, perancangan interface, dan perancangan pengujian. 5. Bab V Hasil dan Pembahasan Bab ini berisi uraian dari peneliti yang menjelaskan system yang dikembangkan. bab ini berisi pemaparan hasil pengujian serta analisa dari penelitian yang telah dibuat 2 6. Bab VI Penutup Bab penutup berisi kesimpulan dan saran. Kesimpulan berisi ringkasan dari keseluruhan hasil penelitian. 4

Saran berisi pendapat dan anjuran yang diberikan oleh peneliti setelah menyelesaikan penelitian BAB II TINJAUAN PUSTAKA Tinjauan pustaka terdiri dari referensi penelitian terdahulu dan referensi teoritis yang diperoleh peneliti. Penjabaran tinjauan referensi sebagai berikut.

1 2.1 Pencapaian terdahulu Peneliti melakukan kajian terhadap jurnal-jurnal yang menjadi acuan dalam 3 penelitian ini. Kajian jurnal terdahulu didasari dengan kemiripan topik yang diambil namun memiliki berbagai perbedaan dengan penelitian yang dilakukan oleh peneliti. 1 2 8 Adapun penelitian ini yaitu melakukan analisis sentimen publik terhadap kebijakan subsidi kendaraan listrik menggunakan arsitektur Long Short Term Memory (LSTM) sebagai salah satu turunan Recurrent Neural Network (RNN). 1 Tabel

2. 1 Pencapaian Terdahulu NO Nama (Tahun) Judul Hasil 1 Mohammed Hafizh Al-Areef dan Kana Saputra (2023) Analisis Sentimen Pengguna Twitter Mengenai Calon Presiden Indonesia Tahun 2024 Menggunakan Algoritma LSTM Hasil dari penelitian ini Hasil uji dan evaluasi performa untuk setiap model yang telah dilatih adalah 82% akurasi, 86% presisi, 92% recall, dan 89% f1-score pada model Ganjar Pranowo. 82% akurasi, 82% presisi, 96% recall, dan 89% f1-score pada model Prabowo Subianto. 87% akurasi, 91% presisi, 95% recall, 93% f1-score

dengan hyperparameter n batch size 32 dan jumlah epoch 50 2 Laina Farsiah, Alim Misbullah, dan Husaini (2022) ANALISIS SENTIMEN MENGGUNAKAN ARSITEKTUR LONG SHORT-TERM MEMORY (LSTM) TERHADAP FENOMENA CITAYAM FASHION WEEK Jumlah training data yang digunakan adalah 80% dari total dataset atau sebanyak 6.043 dan jumlah testing data adalah 20% dari total dataset atau sebanyak 1.511. menghasilkan akurasi sebanyak 88% dengan hyperparameter epoch 50, batch size 128 3 Mitra Wahyudin(2023) Analisis Sentimen pada Data Twitter untuk Memprediksi Hasil Pemilihan Umum 4 M. Adnan Nur Optimasi Berdasarkan 5 skenario pengujian 4 , Nurilmiyanti Wardhani(2022) Normalisasi Kata Pada Data Twitter Untuk Meningkatkan Akurasi Analisis Sentimen (Studi Kasus Respon Masyarakat Terhadap Layanan Teman Bus) yang dilakukan pada penelitian ini, diperoleh tingkat akurasi tertinggi sebesar 0,776 pada skenario ke-3 dan skenario ke-5. Akurasi tersebut didapatkan setelah menerapkan normalisasi kata dengan ratio levenshtein distance untuk koreksi ejaan kata sebesar 0,75 dan 0,9 serta min-count word2vec 5 dan 10. Namun dilihat dari akurasi terendah dari 10 kali percobaan, skenario ke-3 memperoleh akurasi terendah dibandingkan skenario lainnya. Dari hasil tersebut dapat dilihat bahwa akurasi terbaik diperoleh pada skenario ke-5 dengan ratio levenshtein distance sebesar 0,9 dan min-count word2vec sebesar 10. 5 Green Arther Sandag, Jacqueline Waworundeng(2022) Analisis Sentimen Masyarakat Terhadap Exchange Tokocrypto Pada Twitter Menggunakan Metode LSTM Berdasarkan dari hasil dan pembahasan dalam penelitian yang telah dilakukan, dapat diambil kesimpulan bahwa penggunaan model Long Short-Term Memory (LSTM) bekerja dengan baik dalam mengklasifikasikan jumlah tweets tentang tokocrypto yang ada, juga menggunakan Lexicon Sentimen Bahasa Indonesia untuk diambil kata positif, negatif, dan netral. Dari 5000 tweets yang ada mengenai keyword tokocrypto dalam bahasa Indonesia, ditemukan sebanyak 2022 kata positif, 1632 kata negatif, dan 1012 kata netral. 6 Adi Yahyadi dan Fitri Latifah(2022) ANALISIS SENTIMEN TWITTER TERHADAP KEBIJAKAN PPKM DI TENGAH

PANDEMI Pada penelitian ini menghasilkan respon masyarakat terhadap kebijakan PPKM adalah Positif sebesar 27.8%, Netral sebesar 15,7% dan Negatif 5,5%. dalam penelitian ini tingkat akurasi sebesar 70% dengan hyperparameter epoch 10 dan batchsize 128 5 COVID-19 MENGGUNAKAN MODE LSTM 7 Sio Jurnalis Pipin, Heru Kurniawan(2022) Analisis Sentimen Kebijakan MBKM Berdasarkan Opini Masyarakat di Twitter Menggunakan LSTM Pada penelitian ini Model LSTMs yang telah dilatih berdasarkan dataset 658 tweet menghasilkan nilai akurasi terbaik di 80,42%. 7 Hasil analisis sentimen program MBKM dari tweet pengguna antara tahun 2020 hingga 2022 dengan menggunakan kata kunci "MBKM," "MSIB," dan "merdeka belajar" menunjukkan dominasi perasaan "bingung" sebesar 39,51%. Perasaan "bingung" menjadi yang paling umum diungkapkan dalam tweet pengguna. Kemudian, perasaan "senang" muncul sebagai yang kedua terbanyak dengan persentase sebesar 16,26%. Perasaan "sedih" juga cukup signifikan, mencapai 15,80% dari total tweet. Perasaan "marah" menduduki peringkat keempat dengan persentase 13,98%, diikuti oleh perasaan "takut" sebesar 7,29%, dan perasaan "terkejut" sebesar 7,14%. Ini adalah gambaran umum tentang bagaimana pengguna menyatakan perasaan mereka terkait dengan program MBKM selama periode tersebut. 8 Angelina Puput Giovani, Ardiansyah, Tuti Haryanti, Laela Kurniawati, Windu Gata (2020) ANALISIS SENTIMEN APLIKASI RUANG GURU DI TWITTER MENGGUNAKAN ALGORITMA KLASIFIKASI Hasil dari pengujian secara keseluruhan algoritma SVM (PSO) memiliki nilai akurasi dan performa yang paling tinggi 78,55% jika dibandingkan dengan NB (PSO) 67,32%, dan K-NN (PSO) 77,21%. 6 9 Styawati1, Nirwana Hendrastuty, Auliya Rahman Isnain, Ari Yanti Rahmadhani(2021) Analisis Sentimen Masyarakat Terhadap Program Kartu Prakerja Pada Twitter Dengan Metode Support Vector Machine hasil yang diperoleh dari pengujian yang dilakukan kernel linear menghasilkan akurasi 98.67%, presision 98%, recall 99%, dan F1- Score 98%, sedangkan kernel RBF menghasilkan akurasi 98.34%, presision 97%, recall 98%, dan F1- Score 98%. 3 Maka dari sisi akurasi kernel linear lebih baik dari pada RBF 10 Yunico Ardian Pradana,

Imam Cholissodin, Diva Kurnianingtyas Analisis Sentimen Pemindahan Ibu Kota Indonesia pada Media Sosial Twitter menggunakan Metode LSTM dan Word2Vec Hasil penelitian menunjukkan bahwa untuk model dengan kernel linear, akurasi mencapai 98,67%, presisi sebesar 98%, recall 99%, dan F1- Score 98%.

Sementara itu, untuk model dengan kernel RBF, akurasi mencapai 98,34%, presisi sebesar 97%, recall 98%, dan F1-Score 98%. Dengan demikian, dapat disimpulkan bahwa sentimen masyarakat dari pengguna Twitter terhadap program kartu prakerja selama masa pandemi cenderung lebih mendekati netral sebesar 98,34%. Hasil penelitian juga menunjukkan bahwa model dengan kernel linear memiliki akurasi yang sedikit lebih tinggi (98,67%) dibandingkan dengan model kernel RBF (98,34%). Oleh karena itu, dari segi akurasi, model dengan kernel linear terbukti lebih akurat daripada model dengan kernel RBF. 11 Chandra Kirana Poetra, Syarifal Fachri Pane, Rd. Nuraini Siti Fatonah (2021) Meningkatkan akurasi Long- Short term memory (LSTM) pada Analisis Sentimen Vaksin Covid- 19 di Twitter Tujuan dari penelitian ini yaitu membangun aplikasi analisis sentimen menggunakan metode LSTM lalu di optimalisasi dengan menggunakan metode Glove, pengumpulan data menggunakan metode scraping.

1 Preproses data yang digunakan yaitu case folding, stopword removal, dengan Glove translate, cleaning, tokenizing,. Hasil uji akurasi sebesar 89% (87% untuk LSTM dan 88% untuk Bi- LSTM). Dalam penelitian ini penerapan metode Glove dapat meningkatkan akurasi sebesar 2% 12 Esther Irawati Setiawan, Adriel Ferdianto, Joan Santoso, Yosi Kristian, Gunawan, Surya Sumpeno, Mauridhi Hery Purnomo.(2020) Analisis Pendapat Masyarakat terhadap Berita Kesehatan Indonesia menggunakan Pemodelan Kalimat berbasis LSTM Berdasarkan eksperimen yang telah dilakukan dan berfokus pada topik kesehatan, model analisis ini berhasil mencapai rata-rata nilai F1 sebesar 71%. Lebih detailnya, model ini berhasil mencapai nilai 69% untuk label "mendukung," 70% untuk label "menentang," dan 74% untuk label "netral

2 " 13 Rosit Sanusi, Femi Dwi Astuti, dan Indra Yatini Buryadi(2021) ANALISIS SENTIMEN PADA TWITTER TERHADAP PROGRAM KARTU PRA KERJA DENGAN

RECURRENT NEURAL NETWORK Penelitian ini bertujuan untuk mengetahui bagaimana pandangan masyarakat terhadap program kartu prakerja di platform media sosial Twitter. Penelitian tersebut menghasilkan dari 4122 dataset, 2460 netral, 689 positif dan 973 negatif dengan akurasi sebesar 99,31%. 14 Florentina Lalapenta Beru (2023) Analisis Sentimen Masyarakat Terhadap Vaksinasi Sinovac di Twitter dengan Algoritma Long Short Term Memory Dari data set yang telah di kumpulkan dari penelitian ini menghasilkan 535 sentimen netral, 231 sentimen tender, 147 sentimen happy, 98 sentimen sad, 58 sentimen scared, 17 sentimen excited, 15 sentimen angry, serta akurasi sebesar 95%. 15 L. Alfat, K. A. Salleh, M. Nasucha, and N. Sentiment Classification of Indonesian Eksperimen dilakukan dengan berbagai penyesuaian untuk mencapai hasil optimal. Ukuran 8 Baharun, Emotion Related to Vaccination Event using LSTM " in Proceedings of the International Conference on Artificial Intelligence and Internet of Things (AllIoT) batch yang digunakan adalah 100, 200, dan 300 untuk mencari akurasi terbaik dan RMSE terendah. Eksperimen dikonfigurasi dengan epoch = 20, dan pembagian data pelatihan: data pengujian = 80:20. Tabel III menampilkan hasil klasifikasi sentimen pada emosi. Walaupun akurasi keseluruhan melebihi 80%, akurasi tertinggi ditemukan pada ukuran batch = 100. Hasil ini berlaku untuk vaksinasi pertama dengan 98,12%, vaksinasi lansia dengan 99,25%, vaksinasi anak dengan 99,62%, dan vaksinasi booster dengan 99,25%. 1 2.2 Tinjauan teoritis Tinjauan teori digunakan sebagai pendukung teori yang selaras dengan topik pada penelitian ini. Tinjauan teori juga digunakan sebagai pedoman dalam pelaksanaan penelitian yang baik. 1.2 4 22 8 Analisis Sentimen Analisis sentimen adalah proses menganalisis tulisan online untuk menentukan nada emosional dari penulisnya. 4 Selain itu, terdapat pengertian analisis sentimen lain, yaitu proses memahami dan mengelompokkan emosi (positif, negatif, dan netral) pada tulisan dengan menggunakan teknik analisis teks (Faiqotul Himma, 2022). Menurut Faiqotul Himma (2022) Analisis sentimen disebut sebagai opinion mining, yang artinya kata-kata yang

berasal dari masyarakat akan digali emosinya dari setiap kata tersebut. **4** Pendapat masyarakat ini tentu memiliki pengaruh yang besar untuk kemajuan perusahaan atau pemerintah dalam untuk evaluasi usaha, membuat kebijakan, atau merencanakan strategi marketing yang tepat. **1** Opini masyarakat sangat penting bagi perusahaan atau organisasi pemerintahan, karena perusahaan atau organisasi pemerintahan selalu ingin memastikan bagaimana pendapat publik tentang produk, kebijakan dan layanan yang telah dibuat. Dalam sistem pemerintahan, analisis sentimen dapat membantu instansi untuk mendapat opini publik tentang kebijakan yang ingin ditetapkan ataupun kebijakan yang sedang dirancang. Salah satu sumber penambahan data yaitu media sosial. Analisis sentimen sebagian besar menggunakan Natural Language Processing (NLP) untuk memahami perbedaan bahasa atau kalimat yang diamati dalam interaksi pengguna. Fondasi analisis sentimen diletakkan oleh pengembang yang merancang algoritma pembelajaran mesin yang mampu mendeteksi konten yang memiliki beragam sentimen (Vijay Kanadee, 2022). **6 14 8**

NLP Natural Language Processing adalah cabang dari Artificial Intelligence yang berhubungan dengan interaksi antara mesin dan manusia menggunakan bahasa natural (Wahyu Kwann, 2021). Tujuan utama NLP adalah untuk menghubungkan dunia bahasa manusia dengan komputasi dan analisis data. Ini melibatkan pemahaman terhadap struktur bahasa, pemrosesan teks, ekstraksi informasi, analisis sentimen, penerjemahan Bahasa dan generasi teks. Dengan bantuan NLP, komputer dapat menganalisis teks atau ucapan manusia, mengenali pola dan makna dalam bahasa manusia. **1** Salah satu penerapan NLP yaitu menganalisis sentimen user media sosial. Analisis sentimen dilakukan dengan mempelajari pola kalimat positif, negatif dan netral yang ada pada komentar user. Pola kalimat tersebut kemudian digunakan untuk mengklasifikasikan komentar baru ke dalam kategori positif, negatif atau netral. NLP memiliki beberapa pendekatan untuk melakukan analisis sentimen yaitu secara lexicon dan machine learning. Untuk membuat sebuah aplikasi NLP yang handal, dibutuhkan kamus atau kosa kata yang handal. Leksikon dalam pemrograman bahasa alami mengacu pada kelompok

kata, karena kata atau frasa dikaitkan dengan fitur tertentu dan dalam gaya tertentu, seperti bagian dari segmentasi ucapan, serta analisis sentimen, dan mereka juga digunakan. sebagai sumber untuk menafsirkan bahasa manusia dengan memberikan informasi khusus tentang makna dan kegunaan (Mohammed Bakrey, 2022). 3.2 **27** **8 Visual Studio Code**

Visual Studio Code , sering disingkat VS Code , adalah editor kode sumber yang dikembangkan oleh Microsoft. Berikut adalah beberapa poin utama mengenai Visual Studio Code menyediakan berbagai fitur yang memudahkan pengembangan perangkat lunak Python, seperti:

- ☒ Cross-Platform: VS Code dapat diinstal dan digunakan di berbagai sistem operasi, termasuk Windows, macOS , dan Linux .
- ☒ Ringan dan Cepat: Dibandingkan dengan Integrated Development Environments (IDEs) yang lebih berat, VS Code lebih ringan dan cepat, sehingga dapat digunakan di berbagai spesifikasi perangkat keras.
- ☒ Ekstensi dan Kustomisasi: Salah satu kekuatan utama VS Code adalah kemampuannya untuk diperluas melalui ekstensi. Pengguna dapat menambahkan fitur tambahan sesuai kebutuhan, seperti 10 dukungan bahasa tambahan, alat pengembangan, atau integrasi dengan sistem kontrol versi seperti Git.
- ☒ Fitur Canggih: Meskipun ringan, VS Code dilengkapi dengan berbagai fitur canggih seperti auto-completion, debugging, navigasi kode, dan integrasi dengan sistem kontrol. Ini menjadikannya alat yang sangat fleksibel dan powerful untuk pengembangan perangkat lunak.
- ☒ Terminal Terintegrasi: VS Code menyediakan terminal bawaan yang memungkinkan pengembang menjalankan perintah shell langsung dari editor tanpa harus beralih ke aplikasi terminal terpisah.
- ☒ Pengembangan Web dan Cloud : VS Code sangat populer di kalangan pengembang web dan aplikasi berbasis cloud , berkat integrasinya dengan alat-alat pengembangan web dan cloud seperti Docker dan Azure. Secara keseluruhan, Visual Studio Code adalah alat yang sangat berguna untuk pengembangan perangkat lunak berkat fitur-fiturnya yang lengkap, fleksibilitas, dan kemudahan penggunaan.

4.2.8 Media sosial Media sosial adalah platform atau layanan yang

memungkinkan orang untuk berinteraksi, berbagi konten, dan berkomunikasi secara online. Ini memungkinkan pengguna untuk membuat profil pribadi atau halaman, mengirim pesan, membagikan foto, video, dan pemikiran mereka dengan teman-teman, keluarga, atau bahkan dengan orang yang mereka tidak kenal secara langsung. Contoh media sosial termasuk Facebook, Twitter, Instagram, TikTok, LinkedIn, dan banyak platform lainnya. Media sosial telah mengubah cara orang berinteraksi dan berkomunikasi, serta memiliki dampak besar pada budaya dan komunikasi dalam era digital.

1 Selama bertahun-tahun, sistem media sosial menyediakan platform yang baik untuk memfasilitasi dan memungkinkan partisipasi pengguna, keterlibatan, dan komunitas, yang telah menghasilkan budaya partisipatif yang baru (Liu, 2020) 5.2

8 Python Python adalah bahasa pemrograman tingkat tinggi yang populer dan sering digunakan dalam pengembangan perangkat lunak, pengembangan web, analisis data, kecerdasan buatan, pemrograman skrip, dan banyak aplikasi lainnya. 23 11 Dikembangkan pertama kali oleh Guido van Rossum pada akhir 1980-an, Python dirancang dengan fokus pada keterbacaan kode dan kemudahan penggunaan. 1 Bahasa pemrograman ini memiliki library yang beragam dan komunitas yang aktif dalam membangun library yang dapat diakses secara bebas. Berikut merupakan beberapa library yang digunakan dalam penelitian ini. ❏ Natural Language Tool Kit (NLTK) NLTK (Natural Language Toolkit) adalah sebuah perpustakaan (library) dalam bahasa pemrograman Python yang dirancang khusus untuk memproses dan menganalisis teks dalam bahasa alami. NLTK menyediakan berbagai alat dan sumber daya yang membantu dalam tugas-tugas seperti pemrosesan teks, analisis teks, pemrosesan bahasa alami, dan tugas-tugas terkait lainnya. Dengan NLTK, peneliti dapat melakukan berbagai hal, seperti pemisahan kata (tokenization), pemenggalan kata (stemming), pengenalan entitas, analisis sentimen, pemberian label kata, pembuatan model bahasa, dan banyak lagi. 1 NLTK juga dilengkapi dengan berbagai sumber daya seperti korpus teks dan kamus yang dapat membantu dalam penelitian dan pengembangan dalam bidang pemrosesan bahasa alami.

PySastrawi PySastrawi adalah library yang dibuat khusus untuk melakukan pengolahan kata atau kalimat Bahasa Indonesia di bidang NLP (Natural Language Processing). PySastrawi merupakan adaptasi library Sastrawi ke dalam Bahasa pemrograman python yang awalnya hanya mendukung Bahasa pemrograman PHP (Robbani, 2018). PySastrawi digunakan untuk menutupi kekurangan library NLTK yang kurang mendukung pengolahan kata atau kalimat bahasa Indonesia. PySastrawi membantu penggunanya untuk mengubah setiap kata dalam kalimat bahasa Indonesia menjadi kata dasar. Sastrawi juga mampu melakukan stopwords removal atau menghilangkan kata yang tidak memiliki arti pada suatu kalimat bahasa Indonesia. Reguler Expression Regular expression atau Regex adalah sebuah deret karakter yang merepresentasikan sebuah pola yang terdapat pada tipe data string (Khan, R., 2020).

1 9 Regex digunakan untuk pencocokan pola yang terdapat dalam tipe data string untuk dilakukan perubahan atau manipulasi terhadap string tersebut. 1 Python menyediakan fungsi dasar regex di dalam modul re. 12 Regex digunakan untuk melakukan pemrosesan dan manipulasi teks seperti search, replace, dan sub. Dalam regex, setiap karakter memiliki arti khusus yang dapat merepresentasikan karakter atau kelas tertentu. Contohnya dalam pola '[0-9]+' menggambarkan sebuah urutan angka yang dapat muncul dalam sebuah string. Pola tersebut dapat digunakan untuk mengekstrak atau mencari semua karakter yang termasuk dalam pola [0-9]+. SKLearn Scikit-Learn adalah package yang menyediakan versi efisien dari banyak algoritma umum. Scikit-Learn dikategorikan sebagai API sederhana yang bersih dan seragam, dan juga sangat berguna (Hauck, 2014). Scikit-Learn memiliki fitur cross-validation yaitu metode untuk melakukan pemeriksaan keakuratan model dan feature extraction yang berguna untuk melakukan ekstraksi gambar dan teks. Salah satu fitur banyak digunakan oleh peneliti yaitu pembuatan data latih dan data uji. Fitur lainnya yang akan digunakan pada penelitian ini yaitu accuracy_score. Fungsi accuracy_score digunakan untuk melihat akurasi dari model yang telah dibuat.

6.2.8 LSTM Long short term memory atau biasa disebut LSTM

adalah jenis spesial dari RNN yang memiliki kemampuan untuk belajar dengan jangka Panjang. LSTM pertama kali dikenalkan oleh Hochreiter & Schmidhuber pada tahun 1997. LSTM diciptakan untuk mengatasi masalah dependensi jangka Panjang yang terjadi pada Recurrent Neural Network . 13 Gambar 2. 1

Arsitektur LSTM
Arsitektur LSTM terdiri dari beberapa komponen utama, Berikut adalah komponen utama dari LSTM: 1. Cell State (Ct): Ini adalah saluran informasi yang mengalir melalui sel LSTM. Ini digunakan untuk mengirim informasi jangka panjang ke sel berikutnya. 2. Hidden State (Ht): Ini adalah versi yang terfilter dari Cell State dan berfungsi untuk mengirimkan informasi jangka pendek ke sel berikutnya dalam deret waktu. 3. Forget Gate (f): Ini adalah lapisan yang bertanggung jawab untuk memutuskan informasi apa yang akan dihapus atau dilupakan dari Cell State. Ini mengambil Hidden State sebelumnya dan input saat ini sebagai masukan dan menghasilkan angka antara 0 dan 1 untuk setiap elemen dalam Cell State, di mana 0 berarti "lupakan" dan 1 berarti "ingat". 4. Input Gate (i): Input Gate mengatur berapa banyak informasi yang akan ditambahkan ke Cell State. 8 19 Ini juga memutuskan nilai mana yang akan diperbarui. Ini terdiri dari lapisan sigmoid yang menentukan potensi perubahan dan lapisan tanh yang menghasilkan vektor perubahan yang mungkin. 5. Update Cell State (Ct_tilde): Ini adalah langkah di mana informasi baru diperbarui ke Cell State berdasarkan Input Gate dan nilai yang dihasilkan oleh lapisan tanh. 6. Output Gate (o): Ini mengatur seberapa banyak informasi yang akan dikeluarkan ke Hidden State. Ini memproses Hidden State sebelumnya dan input saat ini untuk menghasilkan angka antara 0 dan 1. LSTM memiliki kemampuan untuk mengelola informasi dalam kondisi sel melalui apa yang disebut sebagai gate atau gerbang. Gerbang adalah mekanisme yang memungkinkan pengaturan informasi yang masuk dan keluar serta terdiri dari tiga komponen utama yaitu : lapisan sigmoid dalam jaringan saraf, lapisan hyperbolic tangent (tanh), dan operasi perkalian searah. Lapisan tanh atau aktivasi tanh, serupa dengan sigmoid dalam hal bentuk S,

kelanjutan, dan diferensiasi, namun memiliki rentang nilai antara -1 dan 1, berbeda dengan sigmoid yang memiliki rentang antara dan 1 (Géron, 2022). Tanh digunakan dalam gerbang LSTM untuk mengatasi masalah vanishing gradient dengan memperluas rentang nilai dari -1 hingga 1, dapat dilihat grafik fungsi tanh dalam Gambar 2.2, yang menunjukkan nilai rentangnya dari -1 hingga 1. 14 Lapisan sigmoid, juga dikenal sebagai fungsi sigmoid, adalah suatu fungsi yang mengubah nilai input menjadi nilai di antara dan 1, tetapi tidak pernah mencapai nilai-nilai tersebut secara tepat. 1 Sigmoid digunakan untuk memperkenalkan non-linearitas ke dalam model sehingga model dapat mempelajari hubungan yang lebih kompleks. Grafik fungsi sigmoid, yang ditunjukkan dalam Gambar 2.3, menggambarkan karakteristik bahwa fungsi ini selalu berada di antara 0 dan 1 tanpa pernah mencapai nilai tepat 0 atau 1. Hasil aktivasi sigmoid menggambarkan komponen yang memutuskan apakah informasi akan diteruskan atau tidak. 1 Nilai mendekati menunjukkan bahwa informasi tidak akan diteruskan, sementara nilai mendekati 1 menunjukkan bahwa semuanya akan diteruskan. LSTM memiliki tiga gerbang yang berfungsi untuk menjaga dan mengendalikan kondisi sel. Gambar 2.4 Gambar 2.4, yang ditandai dengan warna merah, menunjukkan langkah awal dalam modul LSTM, yaitu menentukan informasi mana yang akan diteruskan dari kondisi sel sebelumnya. 1 44 Keputusan ini diambil oleh lapisan sigmoid yang disebut 1 17 lapisan gerbang lupakan" atau "forget gate layer 1 "

Lapisan forget 15 Gambar 2. 2 Grafik Tanh Gambar 2. 3 Grafik Fungsi Sigmoid Gambar 2. 4 Forget Gate LSTM gate ini menerima input berupa h_{t-1} dan x_t dengan hasil keluaran berupa nilai antara dan 1 untuk kondisi sel C_{t-1} . Nilai 1 mengindikasikan bahwa informasi akan disimpan sepenuhnya, sementara nilai berarti informasi tersebut tidak akan disimpa 1 . Persamaan dari forget gate adalah sebagai berikut: $f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$ Dimana f_t adalah forget gate, σ adalah fungsi sigmoid yang digunakan untuk menghitung f_t , W_f adalah berat dari forget gate, x_t adalah input pada timestep sekarang, h_{t-1}

adalah vektor hidden layer dari timestep sebelumnya, bf adalah nilai bias dari ft . Gambar 2.5 menunjukkan langkah untuk menentukan informasi terbaru yang ingin disimpan dalam kondisi sel LSTM. Langkah ini melibatkan dua komponen utama, yaitu lapisan sigmoid (sering disebut sebagai "input gate layer") dan lapisan tanh. Lapisan sigmoid, atau "lapisan gerbang masukan," berfungsi untuk memutuskan nilai mana yang akan diubah atau diperbarui. Selanjutnya, lapisan tanh menciptakan vektor kandidat nilai baru, yaitu Ct , yang dapat ditambahkan ke kondisi sel. Langkah selanjutnya adalah menggabungkan hasil dari lapisan sigmoid dan lapisan tanh untuk memperbarui kondisi sel. Persamaan dari lapisan input (input gate layer) dapat dinyatakan sebagai berikut: $it = \sigma(Wi[ht-1, xt] + bi)$ $\sigma = \tanh(Wc[ht-1, xt] + bc)$ Dimana it adalah nilai input gate pada timestamp sekarang, σ adalah lapisan sigmoid, i adalah berat untuk input gate, $ht-1$ adalah output dari timestamp sebelumnya, xt adalah input dari timestamp sekarang, b adalah nilai bias untuk input gate, Ct kandidat nilai baru dari timestamp sekarang, \tanh adalah lapisan tanh, Wc adalah berat untuk kandidat Ct , bc adalah nilai bias untuk Ct .

Gambar 2.5 Input Gate LSTM Gambar 2.6 menggambarkan langkah berikutnya yaitu memperbarui kondisi sel lama atau $Ct-1$ menjadi kondisi sel baru dengan notasi Ct . Adapun caranya yaitu melakukan perkalian kondisi lama dengan ft . hal ini dilakukan untuk melupakan informasi yang telah ditentukan sebelumnya. Kemudian menjumlahkan hasil perkalian it dengan Ct . Berikut merupakan persamaan untuk memperbarui kondisi sel. $Ct = ft \times Ct-1 + it \times Ct$ Dimana Ct adalah kondisi sel pada timestamp t , ft adalah nilai dari gerbang lupa pada timestamp t , $Ct-1$ adalah kondisi sel pada timestamp $t - 1$, it adalah masukan pada timestamp t . Gambar 2.7 menggambarkan langkah terakhir yaitu menentukan luaran. Luaran ini didasarkan dari nilai

kondisi sel yang telah dilakukan filtrasi. Adapun caranya yaitu melewati lapisan sigmoid dan lapisan tanh. Lapisan sigmoid berfungsi untuk menentukan bagian kondisi sel yang ingin dijadikan luaran. Kemudian menempatkan kondisi sel ke dalam lapisan tanh untuk mendapat nilai antara -1 hingga 1 dan melakukan perkalian dengan luaran yang telah melalui lapisan sigmoid. Adapun persamaan yang digunakan pada Langkah ini yaitu sebagai berikut. $ot = \sigma(Wo[ht-1, xt] + bo)$ $ht = ot * \tanh(Ct)$ Dimana ot adalah luaran pada timestamp t , σ adalah fungsi aktivasi sigmoid, W adalah nilai berat pada gerbang luaran, $ht-1$ adalah vektor hidden layer pada timestamp $t-1$, xt adalah input dari timestamp t , o adalah nilai bias dari gerbang.

Gambar 2. 6 Kondisi Sel LSTM Gambar 2. 7 Kondisi Sel LSTM luaran, ht adalah vektor hidden layer pada timestamp t , \tanh adalah fungsi aktivasi hyperbolic tangent, dan Ct adalah kondisi sel baru pada timestamp t .

2.2.7 Flask Flask adalah sebuah web framework mikro untuk bahasa pemrograman Python . Flask dirancang untuk mempermudah pengembangan aplikasi web tanpa memerlukan banyak persiapan dan konfigurasi. Flask, dengan desainnya yang fleksibel, memungkinkan pengembang untuk dengan cepat mengembangkan aplikasi web dengan struktur yang sederhana, di mana framework ini menawarkan fitur-fitur penting seperti sistem routing URL , selain itu Flask juga mendukung pengujian unit dan integrasi secara menyeluruh, sehingga memudahkan pengembang untuk memastikan aplikasi mereka bebas dari bug, semua ini didukung oleh dokumentasi yang sangat baik dan komunitas pengguna yang aktif, menjadikan Flask pilihan yang populer bagi pengembang yang ingin menggabungkan kemudahan penggunaan dengan kemampuan untuk membangun aplikasi web yang skalabel dan dapat disesuaikan.

2.2.8 Black box Metode pengujian " Black Box " adalah pendekatan yang berfokus pada fungsi eksternal yang telah dikembangkan. Dalam metode ini, penelitian dilakukan untuk menemukan kesalahan yang mungkin muncul dalam fungsi, basis data, struktur data, akses data, desain antarmuka, kinerja, dan aspek lainnya.

Pengujian Black Box memeriksa perangkat lunak tanpa memperhatikan rincian internalnya. Proses awal dalam pengujian Black Box melibatkan menetapkan input, yang kemudian diuji untuk mengidentifikasi kesalahan dalam sistem (Priyaungga et al., 2020). Dalam pengujian sistem ini, semua fitur diuji secara menyeluruh untuk memastikan bahwa proses yang dijalankan menghasilkan keluaran atau output sesuai dengan yang telah direncanakan.

2.2.9 White Box Metode pengujian "White Box" adalah pendekatan yang menekankan pada analisis struktur penulisan kode perangkat lunak. Tujuan dari pengujian White Box adalah untuk memastikan bahwa setiap jalur independen dalam struktur internal diperiksa setidaknya satu kali (Setiawan et al., 2020). Metode ini digunakan untuk mengevaluasi apakah baris-baris kode berfungsi sesuai dengan rencana yang telah ditetapkan (Ramdani, 2018). Pengujian White Box melibatkan analisis keluaran dari kode program berdasarkan input yang diberikan. Metode White Box memeriksa keseluruhan logika pemrograman untuk meminimalkan kemungkinan terjadinya kesalahan dalam hasil keluaran.

18 BAB III TAHAPAN PELAKSANAAN Dalam penelitian ini tahap pelaksanaan terdiri dari langkah-langkah pelaksanaan, metode pengembangan sistem dan metode pengujian.

3.1 Langkah-langkah Pelaksanaan Sebuah flowchart langkah pelaksanaan penelitian adalah suatu alat visual yang digunakan untuk menggambarkan penjelasan dan alur penelitian yang akan dilakukan oleh seorang peneliti dalam rangka menyelesaikan tugas akhirnya. Flowchart ini terdiri dari serangkaian tahapan yang diikuti peneliti dalam menjalankan penelitian, dan membantu dalam memahami dan mengkomunikasikan langkah-langkah yang diperlukan dalam penelitian tersebut.

19 1. Perumusan masalah Tahap pertama dalam penelitian ini adalah mengidentifikasi masalah. Masalah ini kemudian diformulasikan untuk memfasilitasi pencarian solusi dalam penelitian ini. Peneliti melakukan observasi terkait isu subsidi kendaraan listrik dan menentukan sumber data yang akan digunakan.

2. Kajian literatur Dalam tahap kajian literatur penelitian ini, peneliti melakukan studi terhadap sumber bacaan yang telah dipublikasikan yang terkait dengan analisis sentimen,

arsitektur LSTM, dan isu subsidi kendaraan listrik. Sumber-sumber yang digunakan untuk mendukung penelitian ini meliputi jurnal, buku, laporan studi, penelitian sebelumnya, serta dokumen-dokumen lain yang relevan. 3. Pengumpulan data dan pengolahan data Penelitian ini menggunakan data opini masyarakat terkait subsidi kendaraan listrik. Data sampel yang digunakan adalah data pada akhir tahun 2022 dengan rentang waktu dari 1 Desember hingga 1 Agustus 2023. Sumber data yang dikumpulkan berasal dari media sosial Twitter, dan proses pengumpulan data dilakukan dengan menggunakan metode scraping API Twitter. Metode scraping memungkinkan pengumpulan data dalam jumlah besar dalam waktu yang relatif singkat dibandingkan dengan pengumpulan data manual. Data yang diambil merupakan tweet yang mengandung kata kunci tertentu dan telah difilter untuk bahasa Indonesia. Setelah pengumpulan data, data tersebut kemudian masuk ke tahap pra-proses data. Tujuan dari tahap ini adalah membersihkan data yang awalnya kotor menjadi data yang bersih dan siap untuk digunakan dalam analisis penelitian. 4. Perancangan dan pengujian model Perancangan model NLP adalah komponen kunci dalam penelitian ini. Model NLP dibangun dengan menggunakan framework TensorFlow 20 Gambar 3. 1 Flowchart Penelitian dan mengadopsi arsitektur LSTM. Model tersebut kemudian disimpan dalam format dengan ekstensi h5. Peran utama dari model ini adalah untuk mengatasi permasalahan yang diajukan dalam penelitian ini, yaitu menentukan opini publik berdasarkan masukan yang diberikan oleh pengguna. Model NLP ini bertugas untuk menganalisis dan menginterpretasikan opini-opini tersebut. 5. Implementasi

21 model ke dalam sistem Tahap implementasi ini melibatkan pembuatan aplikasi sesuai dengan rancangan yang telah disusun. Framework yang digunakan dalam pengembangan aplikasi web ini adalah Flask, dan bahasa pemrograman yang digunakan adalah Python. Implementasi aplikasi ini terdiri dari dua sisi utama, yaitu antarmuka pengguna (front-end) dan logika pengolahan data (back-end). Sisi front-end bertanggung jawab atas tampilan aplikasi yang dilihat oleh pengguna. Ini termasuk tampilan

grafis, elemen antarmuka pengguna, serta cara pengguna berinteraksi dengan aplikasi. Sisi back-end mengacu pada logika pengolahan data dan fungsionalitas aplikasi. Di sini, model NLP yang telah dilatih dan tokenizer yang diperoleh dari dataset digunakan untuk menganalisis dan mengolah data yang masuk. Back-end juga bertanggung jawab untuk memberikan hasil analisis kepada pengguna melalui antarmuka pengguna. Kombinasi antara front-end dan back-end menciptakan aplikasi lengkap yang memungkinkan pengguna untuk berinteraksi dengan model NLP dan mendapatkan hasil analisis terkait dengan opini publik berdasarkan masukan yang mereka berikan.

6. Pengujian sistem Sistem yang ¹² ⁴² telah dikembangkan akan mengalami serangkaian pengujian untuk meminimalisasi kesalahan program atau bug. Pengujian sistem akan menggunakan metode black box dan white box. Metode black box dipilih dengan tujuan untuk memastikan bahwa aplikasi bekerja dengan baik dan sesuai dengan fungsinya. Dalam pengujian black box ³⁵ x, fokus utama adalah pada input dan output aplikasi, tanpa mempertimbangkan rincian internal kode. Hal ini membantu memastikan bahwa aplikasi dapat digunakan sesuai dengan harapan pengguna. Metode white box dipilih untuk memastikan bahwa kode yang telah ditulis berjalan dengan baik. Dalam pengujian white box, fokus utama adalah pada struktur internal kode, serta bagaimana kode tersebut dieksekusi. Hal ini membantu memastikan bahwa setiap bagian dari kode berfungsi dengan benar dan efisien. Kombinasi antara pengujian black box dan white box membantu memastikan kualitas dan keandalan sistem yang dikembangkan, serta meminimalisasi kemungkinan kesalahan atau bug dalam aplikasi.

3.2 Metode Pengembangan Perangkat Lunak

Pengembangan perangkat lunak memiliki beberapa metode. Peneliti menggunakan metode purwarupa dalam mengembangkan perangkat lunak. Metode purwarupa (prototype) adalah pendekatan dalam pengembangan produk atau perangkat lunak yang melibatkan pembuatan model awal atau versi sederhana dari sistem yang akan dikembangkan. Tujuan dari metode ini adalah untuk menguji ide, konsep, atau desain sebelum implementasi penuh. Berikut adalah beberapa aspek dari metode purwarupa:

- ☒ Pembuat

n Model Awal: Metode purwarupa melibatkan pembuatan versi awal dari produk atau sistem, yang dikenal sebagai purwarupa. Model ini bisa berupa sketsa, mockup, atau versi fungsional sederhana dari sistem akhir. ❑ Uji Coba dan Umpan Balik: Purwarupa digunakan untuk mengumpulkan umpan balik dari pengguna atau stakeholder. Dengan menguji model awal ini, pengembang dapat mengidentifikasi masalah, kebutuhan, dan preferensi pengguna sebelum mengembangkan sistem akhir. ❑ Iterasi dan Penyempurnaan: Proses pembuatan purwarupa biasanya bersifat iteratif. Berdasarkan umpan balik yang diterima, purwarupa diperbaiki dan disempurnakan dalam beberapa iterasi hingga mencapai desain yang diinginkan. ❑ Pengurangan Risiko: Dengan mengidentifikasi dan mengatasi masalah pada tahap awal, metode purwarupa membantu mengurangi risiko dan potensi kegagalan dalam pengembangan produk akhir. ❑ Komunikasi: Purwarupa juga berfungsi sebagai alat komunikasi yang efektif antara tim pengembang dan stakeholder, membantu untuk memastikan bahwa kebutuhan dan ekspektasi dipahami dengan jelas. ❑ Penghematan Biaya dan Waktu: Meskipun pembuatan purwarupa memerlukan investasi awal, ini dapat menghemat biaya dan waktu dalam jangka panjang dengan mencegah perubahan besar yang mahal pada tahap akhir pengembangan. Secara keseluruhan, metode purwarupa adalah pendekatan yang berguna untuk merancang dan mengembangkan sistem atau produk dengan lebih efisien, memastikan bahwa solusi akhir lebih sesuai dengan kebutuhan dan harapan pengguna.

22 3.3 Metode Pengumpulan Data

Pada tahap pengumpulan data, penulis mencari dan menyimpan informasi yang dibutuhkan dalam pembangunan sistem. Pengumpulan data dilakukan dengan Teknik scrapping. Pengambilan data menggunakan scraping dengan API adalah metode yang efisien dan terstruktur, di mana aplikasi mengirim permintaan ke endpoint API untuk mengakses dan mengunduh data yang diperlukan dari server, memungkinkan peneliti untuk mendapatkan informasi yang relevan dengan cara yang lebih cepat dan terorganisir dibandingkan dengan scraping tradisional dari halaman web.

3.3 Metode Pengujian

Metode pengujian memang berperan penting dalam memastikan kualitas

dan kinerja sistem atau produk yang dikembangkan. Beberapa manfaat utama dari metode pengujian meliputi: 1. Memastikan Kinerja yang Baik: Metode pengujian membantu memastikan bahwa sistem atau produk berjalan sesuai dengan spesifikasi dan berfungsi dengan baik sesuai dengan tujuannya. 2. Evaluasi Kualitas: Pengujian memberikan kesempatan untuk mengevaluasi kualitas produk atau sistem, yang mencakup aspek-aspek seperti keandalan, kecepatan, keamanan, dan lain-lain. 3. Pendeteksian Kesalahan: Salah satu manfaat utama dari pengujian adalah kemampuannya untuk mendeteksi kesalahan atau bug dalam sistem sebelum produk tersebut dirilis ke pengguna akhir. Ini membantu dalam mengurangi risiko masalah yang mungkin muncul di kemudian hari. 4. Validasi Persyaratan: Metode pengujian membantu memvalidasi bahwa sistem atau produk memenuhi persyaratan dan harapan yang telah ditetapkan. Dalam penelitian ini metode pengujian yang digunakan adalah black box dan white box, yang merupakan pendekatan yang umum digunakan untuk memeriksa fungsionalitas dan integritas perangkat lunak. Kombinasi kedua metode ini dapat membantu memastikan bahwa aplikasi analisis sentimen yang dikembangkan bekerja dengan baik, dan meminimalisasi risiko kesalahan atau bug sebelum aplikasi tersebut dirilis kepada pengguna akhir.

3.4.1 Black Box Pengujian

black box berfokus pada fungsionalitas sistem dari perspektif pengguna atau sisi depan tanpa memeriksa detail kode atau alur data yang ada di dalamnya. Pada pengujian black box, setiap fungsi atau fitur yang ada di antarmuka pengguna (front-end) diuji dengan cara menekan tombol, memberikan input pada formulir, mengunggah berkas, serta mengamati hasil keluaran dan tampilan aplikasi. Tujuannya adalah memastikan bahwa aplikasi berperilaku sesuai dengan harapan pengguna dan dapat digunakan dengan baik. Pengujian black box adalah metode yang sangat berguna untuk memastikan bahwa pengguna akhir akan memiliki pengalaman yang baik dengan aplikasi, dan masalah atau kesalahan yang mungkin muncul dapat diidentifikasi dan diperbaiki sebelum aplikasi dirilis.

3.3.1 White Box Pengujian

white box memang difokuskan pada

pemeriksaan struktur kode dalam tahap perancangan. Tujuan utamanya adalah memastikan bahwa kode berfungsi dengan baik dan bahwa proses input dan output kode bekerja sesuai yang diharapkan. Salah satu teknik yang digunakan adalah statement coverage, Teknik tersebut melibatkan pengujian minimal satu kali terhadap setiap baris kode. Ini berarti setiap baris kode harus dieksekusi selama pengujian untuk memastikan bahwa tidak ada bagian dari kode yang tidak diuji. Dengan cara ini, kesalahan atau potensi bug dalam kode dapat diidentifikasi dan diperbaiki. Pengujian white box ³¹ memiliki peran penting dalam memastikan kualitas kode dan mengurangi risiko kesalahan dalam perangkat lunak. ² Ini membantu memastikan bahwa setiap bagian dari kode berjalan dengan baik dan sesuai dengan perancangan yang telah dibuat. BAB IV PERANCANGAN Pada bab ini menjelaskan secara detail mengenai analisis sistem ² ¹² ¹⁵ rdahulu, spesifikasi sistem kebutuhan baru, perancangan sistem, perancangan tatap muka dan langkah-langkah perancangan sistem. 4.1 Analisis Sistem Terdahulu Analisis sistem terdahul ² bertujuan untuk mengevaluasi cara kerja, kelebihan, kekurangan, dan kebutuhan dari sistem sebelumnya. Pada tahap ini, peneliti dapat mengidentifikasi kebutuhan sistem dengan melakukan pengamatan. Analisis sistem yang telah ada juga membantu memastikan per ⁴⁰ bahan yang akan dilakukan pada sistem yang akan dikembangkan dan menga ² asi kekurangan yang terdapat pada sistem sebelumnya. Pada penelitian yang dilakukan oleh Adi Yahyadi dan Fitri La ifah (2022). Topik yang dibahas pada sistem terdahulu adalah sentimen terhadap kebijakan PPKM di Tengah pandemi covid-19. Metode pengolahan data atau pra-proses data yang digunakan hanya mencakup pembersihan data, case folding, dan tokenisasi. 4.2 Spesifikasi Kebutuhan Sistem Baru Tahap yang dilakukan setelah analisis sistem sebelumnya adalah menetapkan spesifikasi kebutuhan untuk sistem yang akan dikembangkan. Spesifikasi sistem ini penting untuk mencapai tujuan penelitian, yaitu pengembangan aplikasi analisis sentimen menggunakan metode LSTM. Kebutuhan sistem meliputi beberapa komponen yang dijelaskan sebagai berikut. 1. Spesifikasi

Perangkat Lunak (Software) Perangkat lunak adalah elemen krusial dalam proses pengembangan sistem. Dalam penelitian ini, spesifikasi perangkat lunak ditetapkan oleh peneliti untuk mendukung pengembangan sistem. Di bawah ini adalah tabel spesifikasi perangkat lunak yang digunakan.

Tabel 4. 1 Kebutuhan Perangkat Lunak No Kebutuhan Perangkat Lunak 1

Web Browser 2 Visual Studio Code 3 Python 3.9 4 Sistem Operasi

Windows 5 Koneksi Internet Tabel 4.1 berisi spesifikasi perangkat lunak

yang digunakan oleh peneliti dalam pengembangan aplikasi dan model

LSTM. Web browser digunakan untuk menampilkan dan menjalankan aplikasi. Interpreter Pyt

on 3.9 dipilih karena stabil dan telah lama dirilis. Sistem

operasi yang digunakan adalah Windows, yang terhubung ke internet. 2. Spesifikasi Per

angkat Keras (Hardware) Perangkat keras diperlukan untuk

mendukung pengoperasian perangkat lunak dan perangkat keras lainnya.

Berikut ini adalah perangkat keras yang digunakan peneliti dalam

pengembangan aplikasi. Tabel 4. 2 Kebutuhan Perangkat keras No

Perangkat Keras Spesifikasi 1 Processor Intel i 25 2 RAM 8GB Tabel

4.2 mencantumkan perangkat keras yang digunakan untuk menjalankan seluruh

perangkat lunak. Peneliti menggunakan RAM 8GB untuk menghindari lag saat

mengolah dataset. Prosesor yang digunakan adalah Intel Core i5 untuk

melakukan pelatihan model LSTM. Semua data yang terkait dengan aplikasi

dan model disimpan pada HDD. 3. Pembuatan Model Gambar 4.1 adalah flowchart

yang digunakan untuk membuat dan menguji model NLP. Proses

pembuatan model merupakan bagian utama dalam penelitian ini. Tahapan

yang dilakukan meliputi pengumpulan data, pra-proses data, pelabelan data,

pembagian data latih dan uji, tokenisasi kata, pembuatan model, melakukan

prediksi data menggunakan berbagai skenario pengujian, serta evaluasi

model di berbagai skenario pengujian. Pada tahap pengumpulan data,

penulis mencari dan menyimpan informasi yang dibutuhkan dalam pembangunan

sistem. Pengumpulan data dilakukan dengan Teknik 26 Gambar 4. 1

Flowchart Pembuatan Model scrapping. Pengambilan data menggunakan scrapping

dengan API adalah metode yang efisien dan terstruktur, di mana

aplikasi mengirim permintaan ke endpoint API untuk mengakses dan mengunduh data yang diperlukan dari server, memungkinkan peneliti untuk mendapatkan informasi yang relevan dengan cara yang lebih cepat dan terorganisir dibandingkan dengan scraping tradisional dari halaman web. Kata kunci yang digunakan adalah subsidi kendaraan listrik dan kendaraan Listrik kemudian disimpan dalam format xlsx. Praproses data (data preprocessing) adalah serangkaian langkah yang dilakukan untuk mem 7 7 ersiapkan dan membersihkan data mentah sebelum digunakan dalam analisis atau model pembelajaran mesin. Proses ini penting untuk meningkatkan kualitas data sehingga dapat menghasilkan hasil analisis yang lebih akurat dan andal pada tahap selanjutnya. ☒ Case folding adalah proses dalam pengolahan teks yang mengubah semua karakter huruf besar (uppercase) menjadi huruf kecil (lowercase). Tujuan dari casefolding adalah untuk memastikan konsistensi dalam pemrosesan teks, di mana kata-kata atau frasa yang sama tetapi ditulis dengan huruf besar atau kecil dianggap identik. Proses ini penting dalam analisis teks dan pengolahan bahasa alami untuk memastikan bahwa perbandingan string atau pencarian teks tidak dipengaruhi oleh perbedaan kapitalisasi. ☒ Cleaning atau pembersihan dalam konteks pengolahan data adalah proses menghapus atau memodifikasi data yang tidak diperlukan, tidak valid, atau tidak relevan dari dataset. Tujuan utama dari proses cleaning adalah untuk meningkatkan kualitas data dengan menghilangkan noise atau gangguan, sehingga data yang tersedia menjadi lebih konsisten dan dapat diandalkan untuk analisis lebih lanjut. Beberapa langkah yang umum dilakukan dalam proses cleaning data meliputi penghapusan data yang hilang atau tidak lengkap, penanganan nilai yang hilang, deduplikasi data, dan normalisasi format data. Proses ini penting karena data yang bersih dan terstruktur akan memungkinkan pengguna untuk membuat keputusan yang lebih baik dan membangun model yang lebih akurat dalam berbagai aplikasi analisis data dan pembelajaran mesin. 27 Gambar 4. 2 Flowchart Pra-proses Data ☒ Normalisasi kata adalah proses dalam pengolahan tek

s yang bertujuan untuk mengubah variasi kata ke bentuk standar atau dasar yang sama. Tujuan utama dari normalisasi kata adalah untuk mengurangi kompleksitas dan meningkatkan konsistensi dalam analisis teks atau pemodelan bahasa. ✘ Tokenisasi adalah proses dalam pengolahan teks yang mengubah sebuah teks atau dokumen menjadi serangkaian token. Token adalah unit terkecil dari teks yang memiliki makna, seperti kata atau tanda baca. Proses tokenisasi **33** si membagi teks menjadi token-token ini berdasarkan aturan **16** ertentu, seperti spasi antar ka **33** a atau tanda baca sebagai pemisah. Con **16** oh sederhana dari tokenisasi adalah me **33** gub h kalimat "Saya suka makan nasi goreng" menjadi token individu seperti ["Saya", "suka", "makan", "nasi", "goreng"]. Tokenisasi adalah langkah penting dalam pemrosesan bahasa alami dan analisis teks karena memungkinkan komputer untuk memahami dan mengolah teks dalam bentuk yang lebih terstruktur dan dapat dimengerti. ✘ Stopword removal adalah proses dalam pengolahan teks yang bertujuan untuk menghapus kata-kata umum yang tidak memberikan nilai tambah dalam analisis teks atau pemodelan bahasa. Kata-kata ini disebut sebagai "stopwords" karena mereka sering muncul dalam teks namun tidak membawa makna atau informasi yang signifikan untuk pemahaman konten atau tujuan analisis tertentu. Tujuan dari stopwords removal adalah untuk meningkatkan kualitas analisis teks dengan mengurangi noise atau kata-kata yang tidak relevan, sehingga memfokuskan perhatian pada kata-kata kunci atau fitur penting dalam teks. Hal ini dapat meningkatkan kinerja dalam tugas-tugas seperti klasifikasi teks, analisis sentimen, atau pencarian informasi. ✘ Stemming adalah proses dalam pengolahan teks yang menghilangkan infleksi atau afiks dari kata-kata untuk menghasilkan bentuk dasar atau kata dasar yang disebut "stem". Tujuan utama dari stemming adalah untuk mengurangi kata-kata yang berbeda bentuknya menjadi bentuk yang sama sehingga kata-kata tersebut dapat dihitung sebagai satu entitas yang sama dalam analisis teks. Contoh sederhana dari stemming adalah mengubah kata-kata seperti "berjalan", "berjalanlah", dan "berjalan-jalan" menjadi kata dasar "jalan".

☒ Pelabelan data adalah proses memberikan tag atau label pada data untuk mengidentifikasi atau mengkategorikan informasi tertentu yang terkandung dalam data tersebut. Tujuan dari pelabelan data adalah untuk memungkinkan sistem komputer atau model pembelajaran mesin untuk memahami dan memproses informasi dengan lebih baik berdasarkan pada klasifikasi atau atribut yang telah ditentukan sebelumnya. Polaritas yang digunakan pada penelitian ini yaitu positif dan negatif.

4.3 Perancangan Sistem

Pada bagian ini, peneliti menjelaskan proses perancangan sistem untuk aplikasi analisis sentimen menggunakan Unified Modeling Language (UML). Model UML mencakup diagram Use Case, dan diagram Activity yang digunakan untuk mendesain sistem yang akan dikembangkan.

4.3.1 Activity Diagram

Gambar 4.3 menunjukkan diagram aktivitas yang menggambarkan alur kerja sistem. Diagram ini melibatkan tiga entitas utama yang saling terhubung: pengguna, sistem, dan model NLP. Proses dimulai ketika pengguna mengakses aplikasi dan mengunggah file dalam format xlsx. Sistem kemudian menerima file masukan dan melakukan validasi terhadap formatnya. Setelah validasi selesai, sistem membaca konten file dan melakukan praproses data, mirip dengan langkah-langkah pengumpulan dataset pada model NLP. Data yang telah dibersihkan kemudian dianalisis oleh model NLP untuk menghasilkan hasil analisis. Hasil analisis ini disajikan oleh sistem kepada pengguna dalam bentuk statistik, seperti total sentimen yang ditemukan dan sentimen yang paling umum muncul dalam data tersebut.

4.3.2 Flowchart

Gambar 4.3 Activity Diagram menunjukkan flowchart dari sistem analisis sentimen yang dibuat. Tahap pertama dari sistem ini adalah pengguna mengakses sistem dan mengunggah file dalam format xlsx. File yang diunggah kemudian dibaca dan diverifikasi oleh sistem. Jika file sesuai dengan format yang ditentukan, sistem akan melakukan klasifikasi sentimen berdasarkan isi file tersebut. Hasil klasifikasi sentimen diolah oleh sistem untuk menampilkan statistik berupa total sentiment.

4.3.3 Use Case Diagram

Gambar 4.4 Flowchart Penggunaan menunjukkan bahwa pengguna aplikasi analisis sentimen memiliki akses ke berbagai fitur.

Pengguna dapat mengakses aplikasi analisis sentimen, mengunggah file ⁴⁵ ala format yang telah ditentukan ² Selain itu, pengguna juga dapat melihat hasil pemrosesan data yang ditampilkan oleh sistem. 4.3.4 Skenario Use Case

Tabel 4.3 Skenario Use case Unggah File Nama Use Case Unggah File Aktor User Deskripsi User melakukan unggah file ke aplikasi

Tindakan User mengunggah file ke aplikasi sesuai format Tabel 4.1 adalah tabel use case untuk unggah file. Aktor yang dapat menggunakan fitur ini adalah semua pengguna ² yang mengakses aplikasi. Pengunggahan file oleh pengguna harus memenuhi persyaratan yang telah ditentukan. Tabel 4.4 Skenario Use Case

Melihat Data Nama Use Case Melihat Data Aktor User Deskripsi User mampu melihat data yang disajikan sistem Tindakan User melihat data yang disajikan

4.4 Perancangan Antar Muka Perancangan antarmuka adalah proses merancang dan mengembangkan antarmuka pengguna yang mudah digunakan dan efektif untuk sistem yang sedang dibangun. Tujuannya adalah memastikan pengguna dapat berinteraksi dengan sistem dengan mudah dan memahami informasi yang disajikan dengan jelas. Berikut adalah perancangan antarmuka untuk pengguna. 31 Gambar 4.5 Usecase Diagram Aplikasi

Gambar 4.5 merupakan tampilan awal aplikasi, tampilan awal aplikasi menampilkan judul dan panduan untuk menjalankan aplikasi. Gambar 4.6 merupakan tampilan dari menu upload data, Menu unggah data memungkinkan pengguna untuk mengunggah data mereka, yang kemudian akan ditampilkan dalam dashboard tampilan data. Hasil dari tampilan data akan berupa kalimat yang sudah dilabeli sentiment positif dan negative. 32 Gambar 4.6 Rancangan Tampilan awal aplikasi Gambar 4.7 Rancangan Tampilan Menu Upload Gambar 4.7 merupakan rancangan menu TF-IDF pada website ini memungkinkan pengguna untuk menganalisis dan menilai kepentingan kata-kata dalam dokumen. Menu ini juga menampilkan total jumlah tweet keseluruhan ² n, jumlah tweet positif, jumlah tweet negative serta menampilkan kalimat yang sudah dilabeli sentiment positif atau negatif. 4.1 Rancangan Pengujian 4.1.1 Rancangan Pengujian Black Box Tabel 4.3 menggambarkan perencanaan pengujian black box yang mencakup beberapa skenario dan hasil yang

diharapkan. Pengujian black box dirancang untuk mengevaluasi kinerja sistem dan mengurangi kemungkinan bug yang mungkin muncul pada antarmuka. Tabel 4. 5 Skenario Pengujian Black Box No Skenario Pengujian **2** asil yang diharapkan 1 User memilih menu upload lalu memilih opsi choose file setelah itu menekan tombol submit . User dapat memilih menu upload lalu sistem menampilkan halaman upload, Ketika user memilih opsi choose file maka sistem akan menampilkan jendela file explorer untuk memilih file lalu user dapat menekan tombol submit dan sistem menampilkan 33 Gambar 4. 8 Rancangan Tampilan Menu TF IDF data. 2 User memilih menu TF IDF dan menekan tombol TF IDF Sistem menampilkan halaman dan menampilkan hasil berupa jumlah data yang di unggah dan berapa jumlah sentiment positif dan negatif 3 User memilih menu Klasifikasi LSTM Sistem memproses klasifikasi dan menampilkan hasil dari klasifikasi berupa accuracy score , precision score , recall score dan f1 score 4 User memilih menu Tes Model lalu menginput kalimat serta menekan tombol Tes model Sistem memproses k **2** limat yang di masukan user berupa kalimat positif atau negatif lalu menampilkan hasil dari prediksi tersebut 4.5.2 Rancangan Pengujian White Box Tabel 4.4 menggambarkan perencanaan pengujian white box yang terdiri dari beberapa skenario dan hasil yang diharapkan. Pengujian white box dirancang untuk mengevaluasi kesesuaian logika pemrograman berdasarkan input yang diberikan dan memverifikasi bahwa output yang dihasilkan sesuai dengan yang diharapkan. Tabel 4. 6 Skenario Pengujian White Box No Skenario Pengujian Hasil yang di harapkan 1 Aplikasi memuat model dan tokenizer 34 2 Menyamakan tipe huruf menjadi huruf kecil, menghapus username twitter dan hyperlinks 3 Menghilangkan emoji **19** pada kalimat 35 4 Menghil **2** ngkan stopword pada kalimat 5 Mel **3 10 13 37** kukan prediksi yang kemudian dikonversi menjadi label "Positive" atau "Negative" be **3** dasarkan nilai probabilitas. 36 BAB V HASIL DAN PEMBAHASAN Bab ini terdiri dari dua subbab, yaitu hasil dan pembahasan. Subbab hasil menjelaskan hasil pengembangan sistem yang dilakukan dalam penelitian ini. Subbab pembahasan

memberikan penjelasan dan analisis lebih mendalam tentang model NLP serta eng 3 30
 jian sistem. 5.1 Hasil Bagian ini membahas hasil penelitian yang
 telah dilakukan yaitu Pra-proses dan Hasil pemodelan N P. 5.1.1 Pra-proses data
 dan pembuatan model Case Folding Gambar 5.1 merupakan penggalan kod
 e untuk melakukan case folding. Kode mendefinisikan sebuah fungsi
 bernama preprocess_text yang menerima satu parameter text. Baris if
 isinstance(text, str) memeriksa apakah input text adalah sebuah string. Fung 11 29 i
 isinstance digunakan untuk memverifikasi tipe data dari variabel text.
 Jika text adalah string, maka kondisi ini akan benar. 9 11 26 ika text adalah
 string, maka baris lower_case = text.lower() akan mengubah seluruh tek
 s menjadi huruf kecil. Metode .lower() adalah metode string bawaan
 di Python yang mengubah semua karakter huruf besar dalam string
 menjadi huruf kecil. Tabel 5. 1 Hasil case Folding Sebelum Sesudah
 selain pada kasus IKN political will yang cukup tinggi juga terjadi
 pada sektor transisi energi. Jokowi dengan selain pada kasus ikn
 political will yang cukup tinggi juga terjadi pada sektor transisi
 energi jokowi dengan 37 Gambar 5. 1 Kode Penggalan Case Folding
 segala cara membujuk masyarakat untuk beralih ke motor dan mobil
 listrik salah satunya dengan mengeluarkan kebijakan subsidi kendaraan
 listrik. segala cara membujuk masyarakat untuk beralih ke motor dan
 mobil listrik salah satunya dengan mengeluarkan kebijaka 3 subsidi kendaraan
 listrik Cleaning Gambar 5.2 menampilkan potongan kode yang digunaka
 n untuk proses pembersihan data. Proses ini melibatk 3 41 43 n beberapa langkah
 yang memanfaatkan regex untuk menemukan 3 41 pola terentu dan menggantinya
 dengan string yang diinginkan. Hasil dari proses pembersihan data dapat
 dilihat pada Tabel 5.2. Tabel 5. 2 Hasil Cleaning Sebelum Sesudah
 @jokowi Kelilit utang listrik listrik kelilit utang listrik listrik
 masuk pln 38 Gambar 5. 2 Kode Penggalan Cleaning masuk PLN tak
 bisa dijual ... kebijakan kompor listrik...kendaraan listrik...mungkin jika
 3 periode subsidi listrik juga dicabut...keserakahan itu membawa
 kesusahan ! tak bisa dijual kebijakan kompor listrikkendaraan

listrikmungkin jika periode subsidi listrik ju 3 a dicabutkeserakahan itu membawa kesusahan Tabel 5.2 merupakan perbandingan sebelum dan sesudah cleaning data. Adapun perbedaan yang terlihat yaitu hilangnya mentions, tanda baca, dan angka. ✎ Normalisasi Kata Gambar 5.3 menampilkan potongan kode yang digunakan untuk melakukan normalisasi kata. Variabel kamus diisi file CSV yang berisi kamus kata asli dan kata normalisasi input berupa kamus bahasa gaul dan substitusi kata baku dalam bentuk file csv. Lalu teks dipecah menjadi kata-kata individual yang nantinya akan dilak 3 kan iterasi melalui setiap kata dalam teks lalu mencari kata yang sesuai dalam kamus hingga mendapatkan nilai normalisasinya. Hasil normalisasi dapat dilihat pada Tabel 5.3. Tabel 5.3 Hasil Normalisasi Sebelum Sesudah

lucu aq gk prnah pke gas meter lampu pake token non subsidi bbm aq gk prnah pake krena gk suka pake kendaraan mau kluar pesan ajh memasak pun ku pake panci lucu saya tidak pernah pakai gas meter lampu pake token non subsidi bbm saya tidak pernah pake krena tidak suka pake kendaraan mau kluar pesan saja memasak pun 39 Gambar 5.3 Kode Penggalan Normalisasi listri dan kompor listrik loe ajah yg muka ngemis bansos loll jgn lupa bansos yg loe makan jg pajak dri gw ku pake panci listri dan kompor listrik kamu ajah yang muka ngemis bansos loll jangan lupa bansos yang kamu makan juga pajak dri saya ✎ Tokenisasi Menggunakan metode `word_tokenize` dari modul `nltk.tokenize` untuk memecah kalimat menjadi daftar token (kata-kata). Metode ini memisahkan kata-kata dalam kalimat berdasarkan spasi dan tanda baca, mengubahnya menjadi elemen-elemen terpisah dalam daftar. Tabel 5.4 Hasil Tokenisasi Sebelum Sesudah kelilit utang listrik listrik masuk pln tak bisa dijual kebijakan kompor listrikkendaraan listrikmungkin jika periode subsidi listrik juga dicabutkeserakahan itu membawa kesusahan ['kelilit', 'utang', 'listrik', 'listrik', 'masuk', 'pln', 'tak', 'bisa', 'dijual', 'kebijakan', 'kompor', 'listrikkendaraan', 'listrikmungkin', 'jika', 'periode', 'subsidi', 'listrik', 'juga', 'dicabutkeserakahan', 'itu', 'membawa', 'kesusahan'] ✎ Stemming Fungsi `stemming_text` bertujuan untu

k melakukan stemming pada daftar token yang diberikan, dengan pengecualian untuk token 'kendaraan' 40 Gambar 5. 4 Kode Penggalan Tokenisasi Gambar 5. 5 Kode Penggalan Stemming yang tidak di stemming . Stemming adalah proses mengubah kata-kata ke bentuk dasarnya (stem) untuk mengurangi variasi kata-kata yang serupa. Tabel 5. 5 Hasil Stemming Sebelum Sesudah ['ditambah', 'subsidi', 'kendaraan', 'listrik', 'pak'] ['tambah', 'subsidi', 'kendaraan', 'listrik', 'pak'] ✕ Pelabelan Data Fungsi sentiment_analysis_lexicon_indonesia bertujuan untuk menghitung skor sentimen dari sebuah teks berdasarkan kamus sentimen 41 Gambar 5. 6 Kode Penggalan Pelabelan data (lexicon) bahasa Indonesia. Fungsi ini menambahkan nilai positif untuk kata-kata yang terdapat dalam kamus positif dan mengurangi nilai negatif untuk kata-kata yang terdapat dalam kamus negatif. Hasil akhirnya adalah 3 skor sentimen yang menunjukkan seberapa positif atau negatif teks tersebut. 5.1.2 Pra-proses data dan pembuatan model Tabel 5. 6 Hasil Pemodelan Batch Size Epoch Rasio Data Akurasi 32 5 80:20 77,7% 64 5 0:20 70,5% 128 5 80:20 71,3% Tabel 5.6 menggambarkan hasil uji pengaruh hyperparameter terhadap akurasi yang diperoleh. Hyperparameter yang diuji yaitu batch size dan rasio pembagian data. Epoch yang digunakan pada pengujian adalah 5. Epoch menggambarkan setiap 3 sampel dalam dataset pelatihan yang memiliki kesempatan untuk memperbarui parameter model internal satu kali selama satu epoch. Seluruh batch data yang telah melewati model dihitung sebagai satu epoch. Jumlah epoch dapat berapa satu atau lebih. Batch size menggambarkan data pelatihan yang dipotong menjadi bagian-bagian kecil. Batch yang lebih kecil ini dapat dengan mudah dimasukkan ke dalam mode pembelajaran mesin untuk melatihnya. Ukuran batch size memiliki batas maksimal sesuai jumlah data pelatihan dan minimal satu. Pada pengujian hyperparameter batch size dengan ukuran 32 memiliki akurasi rata-rata 77%. Batch size dengan ukuran 64 memiliki akurasi rata-rata 70,5%. Batch size dengan ukuran 128 memiliki akurasi rata-rata sebesar 71,3%. 5.2 Pembahasan Dalam penelitian ini, dua

metode pengujian yang digunakan adalah metode Black Box dan White Box. Berikut ini adalah hasil pengujian dan pembahasan perangkat lunak yang telah dibuat.

5.2.1 Hasil Pengujian Black Box

Black box testing adalah metode pengujian eksternal yang bertujuan memastikan bahwa suatu sistem atau aplikasi dapat memenuhi kebutuhan dan harapan pengguna. Pengujian ini berfokus pada penilaian kualitas semua fitur aplikasi, sehingga penguji dapat memastikan bahwa sistem atau aplikasi berfungsi dengan benar.

Tabel 5. 7 Hasil Pengujian Black Box No Skenario

pengujian Hasil yang diharapkan

- 1 Pengguna klik Choose file Muncul jendela explorer untuk unggah data
Kesimpulan : berhasil memunculkan jendela explorer
- 2 Pengguna klik tombol Submit Menampilkan data yang sudah diberikan label positif dan negatif
Kesimpulan : berhasil menampilkan data
- 3 Pengguna pilih menu TF IDF dan klik tombol Proses TF IDF Menampilkan jumlah data tweet serta jumlah tweet positif dan negatif
- 4 Kesimpulan : Berhasil menampilkan data tweet
- 4 Pengguna pilih menu klasifikasi LSTM dan klik tombol Klasifikasi LSTM Menampilkan hasil klasifikasi
Kesimpulan : berhasil menampilkan nilai hasil klasifikasi
- 5 Pengguna pilih menu Tes Model, menginput teks lalu klik tombol Tes Model Menampilkan hasil prediksi kata
- 4 Kesimpulan : Berhasil menampilkan prediksi kalimat

5.2 Hasil Pengujian White Box

White box testing adalah metode pengujian internal yang memastikan bahwa kode sistem atau aplikasi bebas dari bug dan memenuhi standar kualitas yang ditetapkan. Selain itu, pengujian ini bertujuan untuk memeriksa bagaimana setiap bagian dari sistem atau aplikasi bekerja secara internal menunjukkan beberapa skenario pengujian dengan berisi hasil yang dirapkan, kode program dan hasil pengujian.

Tabel 5. 8 Hasil Pengujian White Box No Hasil yang diharapkan

Kode program Hasil Pengujian

- 45 1 Aplikasi memuat model dan tokenizer Berhasil memuat model dan tokenizer
- 2 Menyamakan tipe huruf menjadi huruf kecil, menghapus username twitter dan hyperlinks Berhasil Menyamakan tipe huruf menjadi huruf kecil, menghapus username twitter dan hyperlinks
- 3

Menghilangkan emoji pada kalimat Berhasil Menghilangkan emoji pada kalimat 46 4 Menghilangkan stopword pada kalimat Berhasil M ngkilangkan stopword pada kalimat 5 Melakukan prediksi yang kemudian dikonversi menjadi label "Positive" atau "Negative" berdasarkan nilai probabilitas. Berhasil Melakukan prediksi yang kemudian dikonvers 5 menjadi label "Positive" atau "Negative" berdasarkan nilai probabilitas. Tabel 5.8 menc 5 kup skenario pengujian white box. Pengujian ini dimulai dengan proses aplikasi memuat model dan tokenizer ke sesi web browser. Hasilnya, aplikasi berhasil memuat model dan tokenizer dengan sukses. Selanjutnya, dilakukan pengujian kode program untuk menghilan 5 kan emoji, menghapus stopword, dan membersihkan teks. Kode berjalan lancar dan menghasilkan output sesuai yang diharapkan. Pengujian berik 14 tnya mencakup kode untuk melakukan word embedding, padding, prediksi, dan mengkodekan output dari bentuk numerik menjadi kata. 47 BAB VI PENU UP Bab ini menjelaskan hasil pengembangan yang telah dilakukan serta pembahasan terkait pengembangan tersebut. Bab ini terdiri dari dua bagian, yaitu kesimpulan dan saran. 6.1 Kesimpulan Pengembangan aplikasi analisis sentimen menggunakan pendekatan pengolahan bahasa alami telah dilakukan. Berikut adalah kesimpulan yang diperoleh dari penelitian ini. 1. Hasil analisis dikembangkan menggunakan konsep statistik dengan nilai- nilai kuantitatif. 2. Pemodelan analisis sentimen dalam studi kasus subsidi kendaraan listrik menggunakan arsitektur LSTM. Model ini mampu melakukan klasifikasi sentimen berdasarkan dataset yang telah dibuat. Akurasi tertinggi yang dicapai oleh model LSTM dengan rasio pembagian data 80:20, ukuran batch 32, dan 5 epoch adalah 77,7%. 3. Sentimen mengenai teknologi dan kebijakan subsidi kendaraan listrik di media sosial Twitter pada awal perilisan desem 13 er 2022 sampai agustus 2023 mendapat sentiment negatif sebanyak 1984 atau 65,9% dan sentimen positif sebanyak 1026 atau 34,1%. 48 6.2 Saran Berdasarkan hasil dan pembahasan yang telah disampaikan pada bab sebelumnya, berikut adalah beberapa saran untuk penelitian selanjutnya. 1. Diharapkan penelitian selanjutnya melakukan filtrasi antara tweet berita dan nonberita untuk meminimalkan

REPORT #22128215

ketidakseimbangan dataset. 2. Sentimen analisis subsidi kendaraan listrik ini dapat dikembangkan dengan membandi



REPORT #22128215

Results

Sources that matched your submitted document.

● IDENTICAL ● CHANGED TEXT

INTERNET SOURCE		
1.	0.97% eprints.upj.ac.id https://eprints.upj.ac.id/id/eprint/6052/9/9.%20BAB%20II.pdf	● ●
INTERNET SOURCE		
2.	0.72% media.neliti.com https://media.neliti.com/media/publications/560286-analisis-sentimen-pada-tw...	●
INTERNET SOURCE		
3.	0.52% ejournal.uniramalang.ac.id https://ejournal.uniramalang.ac.id/index.php/g-tech/article/download/3346/218..	●
INTERNET SOURCE		
4.	0.36% eprints.upj.ac.id https://eprints.upj.ac.id/id/eprint/6098/9/Skripsi%20Final%20-%20Roghib_Bab2..	●
INTERNET SOURCE		
5.	0.31% eprints.upj.ac.id https://eprints.upj.ac.id/id/eprint/6052/12/12.%20BAB%20V.pdf	●
INTERNET SOURCE		
6.	0.31% repository.uinjkt.ac.id https://repository.uinjkt.ac.id/dspace/bitstream/123456789/66253/1/NUR%20AD..	●
INTERNET SOURCE		
7.	0.29% www.academia.edu https://www.academia.edu/120027437/Perbandingan_CNN_dan_Bi_LSTM_pad...	●
INTERNET SOURCE		
8.	0.29% blog.algorit.ma https://blog.algorit.ma/lstm-network/	●
INTERNET SOURCE		
9.	0.28% ejournal.bsi.ac.id https://ejournal.bsi.ac.id/ejurnal/index.php/ijse/article/download/15671/6023	●



REPORT #22128215

INTERNET SOURCE		
10.	0.25% lib.lemhannas.go.id	●
	http://lib.lemhannas.go.id/public/media/catalog/0010-112300000000088/swf/74..	
INTERNET SOURCE		
11.	0.23% repository.uinjkt.ac.id	●
	https://repository.uinjkt.ac.id/dspace/bitstream/123456789/50432/1/AFIFAH%2...	
INTERNET SOURCE		
12.	0.22% e-journal.uajy.ac.id	●
	http://e-journal.uajy.ac.id/30652/2/191124961_Bab%201.pdf	
INTERNET SOURCE		
13.	0.22% e-journal.uajy.ac.id	●
	http://e-journal.uajy.ac.id/4450/6/5EP17948.pdf	
INTERNET SOURCE		
14.	0.18% e-journal.uajy.ac.id	●
	http://e-journal.uajy.ac.id/24407/2/1707091411.pdf	
INTERNET SOURCE		
15.	0.15% eprints.itn.ac.id	●
	http://eprints.itn.ac.id/5128/2/1752029_BAB%201%20-%20muchammad%20nu...	
INTERNET SOURCE		
16.	0.09% bisa.ai	●
	https://bisa.ai/portofolio/detail/MzEwMw	
INTERNET SOURCE		
17.	0.06% digilib.unila.ac.id	●
	http://digilib.unila.ac.id/72040/3/3.%20SKRIPSI%20TANPA%20BAB%20PEMBAH...	
INTERNET SOURCE		
18.	0.03% arxiv.org	●
	https://arxiv.org/pdf/2301.04521	
INTERNET SOURCE		
19.	0.03% eprints.upj.ac.id	●
	https://eprints.upj.ac.id/id/eprint/6052/11/11.%20BAB%20IV.pdf	



REPORT #22128215

● QUOTES

INTERNET SOURCE

1. **12.15%** eprints.upj.ac.id
<https://eprints.upj.ac.id/id/eprint/6052/9/9.%20BAB%20II.pdf>

INTERNET SOURCE

2. **9.38%** eprints.upj.ac.id
<https://eprints.upj.ac.id/id/eprint/6052/11/11.%20BAB%20IV.pdf>

INTERNET SOURCE

3. **3.27%** eprints.upj.ac.id
<https://eprints.upj.ac.id/id/eprint/6052/12/12.%20BAB%20V.pdf>

INTERNET SOURCE

4. **0.75%** majoo.id
<https://majoo.id/solusi/detail/analisis-sentimen-adalah>

INTERNET SOURCE

5. **0.59%** eprints.upj.ac.id
<https://eprints.upj.ac.id/id/eprint/6609/11/BAB%20III.pdf>

INTERNET SOURCE

6. **0.46%** penerbit.gunadarma.ac.id
<https://penerbit.gunadarma.ac.id/wp-content/uploads/2023/09/Pengantar-Meto..>

INTERNET SOURCE

7. **0.4%** bisa.ai
<https://bisa.ai/portofolio/detail/MzA5OA>

INTERNET SOURCE

8. **0.38%** digilib.unila.ac.id
<http://digilib.unila.ac.id/69997/3/3.%20SKRIPSI%20TANPA%20BAB%20PEMBAH...>

INTERNET SOURCE

9. **0.37%** repository.uinjkt.ac.id
<https://repository.uinjkt.ac.id/dspace/bitstream/123456789/76750/1/FAZRIANSY...>

INTERNET SOURCE

10. **0.36%** eprints.upj.ac.id
<https://eprints.upj.ac.id/id/eprint/6060/8/BAB%20I.pdf>

INTERNET SOURCE

11. **0.34%** dasarpemrogramanpython.novalagung.com
<https://dasarpemrogramanpython.novalagung.com/basic/string>



REPORT #22128215

INTERNET SOURCE

12. **0.34%** eprints.upj.ac.id

<https://eprints.upj.ac.id/id/eprint/7526/15/BAB%20IV.pdf>

INTERNET SOURCE

13. **0.3%** repository.uinjkt.ac.id

<https://repository.uinjkt.ac.id/dspace/bitstream/123456789/68556/1/SUCI%20A...>

INTERNET SOURCE

14. **0.28%** algorit.ma

<https://algorit.ma/blog/natural-language-processing/>

INTERNET SOURCE

15. **0.24%** eprints.upj.ac.id

<https://eprints.upj.ac.id/id/eprint/2725/11/BAB%20IV.pdf>

INTERNET SOURCE

16. **0.23%** eprints.upj.ac.id

<https://eprints.upj.ac.id/id/eprint/6097/10/BAB%20III.pdf>

INTERNET SOURCE

17. **0.23%** smarteksistem.com

<https://smarteksistem.com/software-testing/?lang=id>

INTERNET SOURCE

18. **0.22%** repository.umy.ac.id

<http://repository.umy.ac.id/bitstream/handle/123456789/15195/07%20BAB%20...>

INTERNET SOURCE

19. **0.2%** arxiv.org

<https://arxiv.org/pdf/2301.04521>

INTERNET SOURCE

20. **0.2%** lms.onnocenter.or.id

https://lms.onnocenter.or.id/wiki/index.php/Keras:_Difference_Between_a_Bat...

INTERNET SOURCE

21. **0.19%** medium.com

<https://medium.com/@myskill.id/introduction-to-flask-4804d062871f>

INTERNET SOURCE

22. **0.19%** blog.kazee.id

<https://blog.kazee.id/apa-itu-analisis-sentimen/>



REPORT #22128215

INTERNET SOURCE

23. **0.19%** saidalfaraby.github.io

https://saidalfaraby.github.io/pengenalan-pemrograman/9-Pengolahan_string/M

INTERNET SOURCE

24. **0.19%** eprints.unpak.ac.id

<https://eprints.unpak.ac.id/7677/1/Skripsi%20Refansyach%20065119226.pdf>

INTERNET SOURCE

25. **0.18%** repository.stei.ac.id

<http://repository.stei.ac.id/5918/4/BAB%20III%20FINAL%20REVISI.pdf>

INTERNET SOURCE

26. **0.17%** media.neliti.com

<https://media.neliti.com/media/publications/268955-pendeteksi-kesamaan-kat...>

INTERNET SOURCE

27. **0.17%** medium.com

<https://medium.com/@malsky/code-extension-visual-studio-code-d9f49c26de53>

INTERNET SOURCE

28. **0.17%** repo.darmajaya.ac.id

<http://repo.darmajaya.ac.id/10301/9/Bab%202.pdf>

INTERNET SOURCE

29. **0.17%** repository.uinjkt.ac.id

<https://repository.uinjkt.ac.id/dspace/bitstream/123456789/76873/1/FRIDO%20...>

INTERNET SOURCE

30. **0.16%** ejurnal.umri.ac.id

<https://ejurnal.umri.ac.id/index.php/JIK/article/download/2772/1566>

INTERNET SOURCE

31. **0.16%** eskripsi.usm.ac.id

<https://eskripsi.usm.ac.id/files/skripsi/G11A/2017/G.131.17.0092/G.131.17.0092-...>

INTERNET SOURCE

32. **0.15%** repository.teknokrat.ac.id

<http://repository.teknokrat.ac.id/5533/4/b219312248.pdf>

INTERNET SOURCE

33. **0.13%** bisa.ai

<https://bisa.ai/portofolio/detail/MzEwMw>



REPORT #22128215

INTERNET SOURCE

34. **0.13%** dspace.uui.ac.id

<https://dspace.uui.ac.id/bitstream/handle/123456789/16619/05.3%20bab%203.p..>

INTERNET SOURCE

35. **0.12%** digitalsolusigrup.co.id

<https://digitalsolusigrup.co.id/black-box-testing-adalah/>

INTERNET SOURCE

36. **0.12%** jurnal.tau.ac.id

<https://jurnal.tau.ac.id/index.php/snartek/article/download/551/388/2317>

INTERNET SOURCE

37. **0.12%** repository.its.ac.id

https://repository.its.ac.id/60580/1/05211440000118-Undergraduate_Theses.pdf

INTERNET SOURCE

38. **0.12%** kc.umn.ac.id

https://kc.umn.ac.id/14897/5/BAB_III.pdf

INTERNET SOURCE

39. **0.1%** jurnal.upnyk.ac.id

<http://jurnal.upnyk.ac.id/index.php/semnasif/article/viewFile/1016/977>

INTERNET SOURCE

40. **0.09%** sipora.polije.ac.id

<https://sipora.polije.ac.id/16485/3/BAB%201.pdf>

INTERNET SOURCE

41. **0.09%** ejournal.pelitaindonesia.ac.id

<https://ejournal.pelitaindonesia.ac.id/ojs32/index.php/JOISIE/article/download...>

INTERNET SOURCE

42. **0.09%** ejournal.unisbablitar.ac.id

<https://ejournal.unisbablitar.ac.id/index.php/antivirus/article/download/2501/1...>

INTERNET SOURCE

43. **0.08%** ejournal.itn.ac.id

<https://ejournal.itn.ac.id/index.php/jati/article/download/8469/5298/>

INTERNET SOURCE

44. **0.07%** digilib.unila.ac.id

<http://digilib.unila.ac.id/72040/3/3.%20SKRIPSI%20TANPA%20BAB%20PEMBAH...>

REPORT #22128215

INTERNET SOURCE

45. 0.04% eskripsi.usm.ac.id

<https://eskripsi.usm.ac.id/files/skripsi/G11A/2015/G.111.15.0011/G.111.15.0011-...>