

BAB I PENDAHULUAN

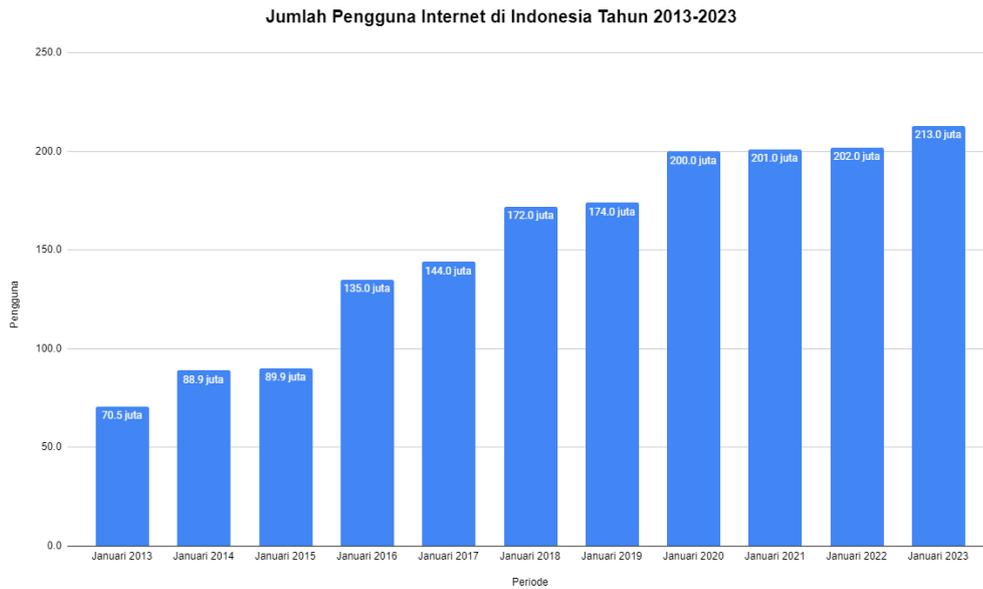
Pada bab pendahuluan akan dijelaskan mengenai gambaran umum terhadap topik penelitian yang dibahas dalam penelitian ini. Bab pendahuluan ini terbagi menjadi beberapa bagian yang terdiri atas latar belakang, perumusan masalah, tujuan penelitian, batasan penelitian, dan manfaat penelitian.

I.1 Latar Belakang

Pemilu Presiden dan Wakil Presiden Indonesia merupakan ajang demokrasi di Indonesia yang dilaksanakan untuk memilih Presiden dan Wakil Presiden Indonesia dalam masa jabatan 5 tahun. Komisi Pemilihan Umum (KPU) mengumumkan bahwa terdapat Daftar Pemilih Tetap (DPT) sebanyak 204.807.222 pemilih dalam ajang Pemilu 2024 (KPU, 2023).

Antusiasme pemilih dalam pemilu di tahun 2024 cukup tinggi. Hal ini didukung oleh survei tim Riset dan Analitik Kompas Gramedia Media bersama Litbang Kompas mengenai partisipasi pemilu 2024 yang menunjukkan bahwa 86,7% responden tertarik untuk berpartisipasi di dalam pemilu 2024 (Rahayu, 2022). Selain itu, pemilu membawa pembahasan pikiran dan pandangan politik yang hangat di masyarakat, baik itu pro maupun kontra (Wenando dkk., 2020). Lebih lanjut, penyampaian pendapat dari masyarakat terkadang mengandung ujaran kebencian yang memberikan efek negatif bagi persepsi masyarakat lain maupun kepada kandidat (Talita & Wiguna, 2019). Dengan munculnya penyampaian pendapat pro dan kontra di masyarakat, para kandidat dalam pemilu perlu untuk mendapatkan wawasan objektif dari perbincangan publik, sehingga para kandidat tersebut dapat memahami dan merespons opini dan kritik yang berasal dari masyarakat.

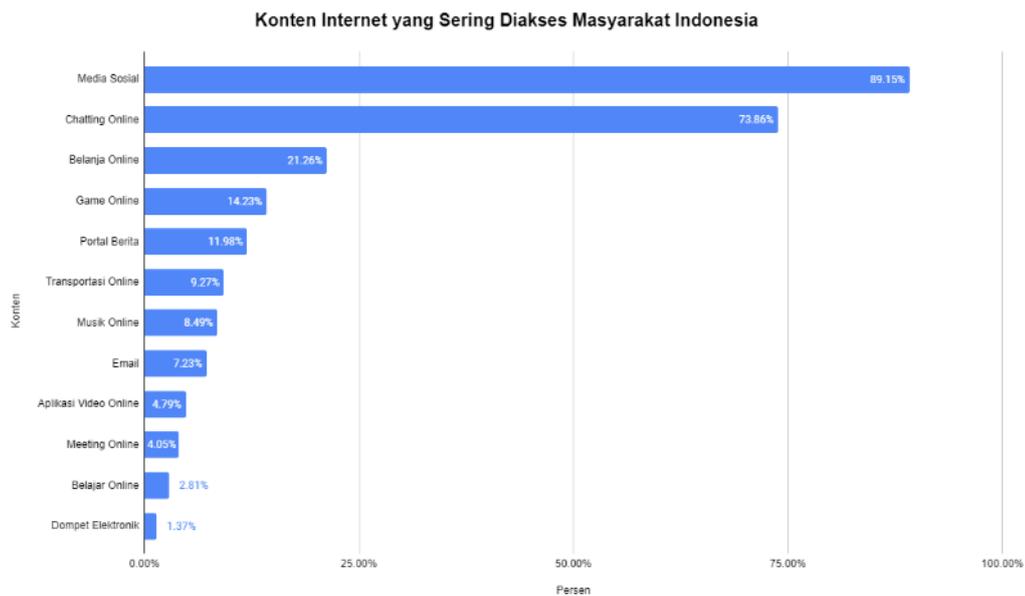
Karena jumlah pemilih yang besar tersebut, pemanfaatan internet dapat digunakan menjadi salah satu metode untuk melakukan kampanye demi mendapatkan dukungan dari masyarakat (Dewi & Widyawati, 2021). Berkembangnya internet di Indonesia yang cukup masif juga membuat pengguna internet di Indonesia semakin meningkat. Hal ini didukung oleh data yang dihimpun oleh *We Are Social* seperti tertera pada Gambar I.1.



Gambar I.1 Jumlah Pengguna Internet di Indonesia (We Are Social, 2023)

Berdasarkan Gambar I.1, dapat dilihat bahwa pada bulan Januari 2023 terdapat 213 juta pengguna internet di Indonesia. Jumlah tersebut naik 5,2% dari periode sebelumnya yaitu di bulan Januari 2022 di mana Indonesia memiliki pengguna internet sebanyak 202 juta pengguna. Besarnya pengguna internet di Indonesia tersebut dapat dimanfaatkan oleh para pasangan calon presiden maupun wakil presiden untuk Pemilu Presiden dan Wakil Presiden Indonesia 2024. Hal tersebut dapat dilakukan dengan memanfaatkan internet sebagai alat kampanye yang efektif dan sarana komunikasi dengan publik yang efektif pula (Dalimunthe, 2017).

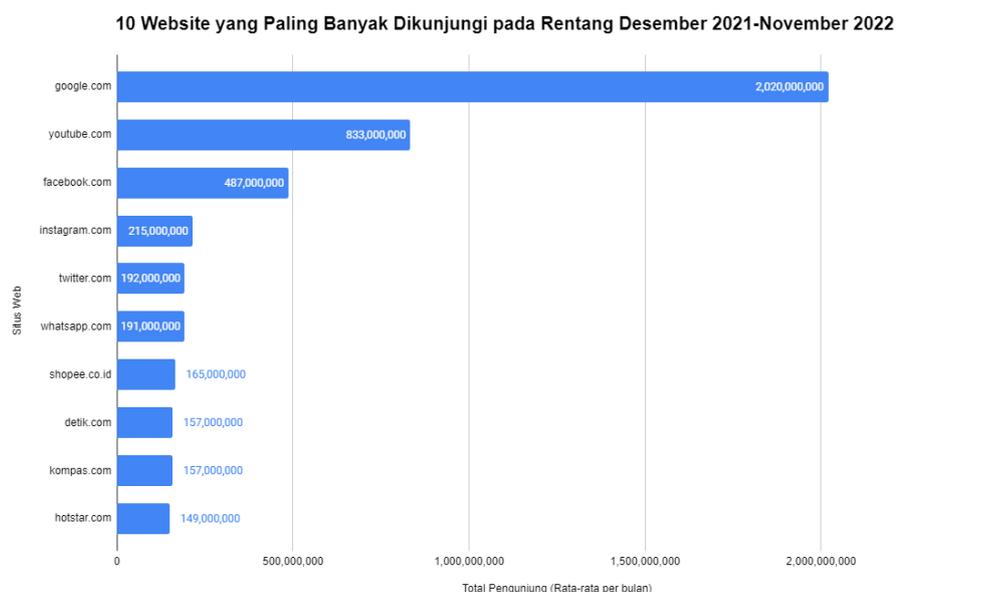
Menyikapi hal tersebut, diperlukan strategi politik untuk dapat menarik suara pemilih. Salah satu media yang dapat digunakan untuk menarik suara yaitu media sosial. Hal ini karena terdapat kecenderungan masyarakat Indonesia dalam mengakses media sosial. Pada Gambar I.2, dapat dilihat hasil dari survei Asosiasi Penyelenggara Jasa Internet Indonesia (APJII) yang menyatakan bahwa pada rentang 2021-2022, 89,15% responden menggunakan internet di Indonesia untuk mengakses media sosial.



Gambar I.2 Konten Internet yang Sering Diakses Masyarakat Indonesia
(Databoks, 2022)

Seiring dengan perkembangan internet, pengguna internet kian bertambah dan generasi muda atau generasi milenial memilih internet sebagai sumber berita utama mereka untuk topik politik (Adinugroho dkk., 2019). Selain itu, penggunaan media sosial sendiri juga dianggap efektif dan penting oleh aktor politik untuk digunakan sebagai media demi penyebaran informasi untuk membentuk opini maupun agenda politik (Perdana & Wildianti, 2018). Salah satu media sosial yang banyak digunakan di Indonesia adalah YouTube. YouTube merupakan platform media sosial yang dimiliki oleh Alphabet Inc. dan dapat digunakan untuk aktivitas mengunggah video maupun menonton video.

Melansir dari data yang dipublikasikan oleh *We Are Social* pada bulan Januari 2023, YouTube berada di peringkat 8 pada *query* pencarian Google sepanjang tahun 2022 karena memiliki perbandingan volume pencarian dengan memiliki indeks pencarian sebesar 27 jika dibandingkan dengan *top query* yang ada. *We Are Social* juga memaparkan lebih lanjut mengenai peringkat situs web yang banyak dikunjungi oleh pengguna internet di Indonesia pada rentang bulan Desember 2021 hingga November 2022, infografis peringkat tersebut disajikan pada Gambar I.3



Gambar I.3 *Website* yang Paling Banyak Dikunjungi pada Rentang Desember 2021- November 2022 (We Are Social, 2023)

Berdasarkan Gambar I.3 dapat dilihat bahwa YouTube menempati peringkat 2 sebagai situs web yang paling banyak dikunjungi oleh pengguna internet Indonesia di rentang yang diberikan dengan total pengunjung rata-rata per bulan sebanyak 833 juta pengunjung. Besarnya pengunjung YouTube juga akan menghasilkan data yang besar pula di platform tersebut. Satu dari sekian banyak data yang dapat dihasilkan di YouTube adalah komentar dari suatu video yang diunggah. Komentar-komentar di YouTube tersebut kemudian dapat digunakan sebagai parameter untuk menilai dan mengevaluasi kanal yang ada di YouTube (Munthe dkk., 2021). Dengan ditayangkannya video debat kepresidenan pada Pemilu Presiden dan Wakil Presiden Indonesia 2024 secara langsung lewat YouTube, ruang diskusi bagi publik akan terbuka lebar melalui kolom komentar. Komentar yang disampaikan terdiri atas pendapat yang pro dan kontra terhadap masing-masing pasangan kandidat. Berdasarkan komentar-komentar yang ada pada video tersebut, dapat dilakukan sebuah analisis berupa analisis sentimen.

Menurut Oktaviana dkk. (2022), analisis sentimen merupakan sebuah studi untuk menganalisis suatu teks agar dapat mengetahui suatu pendapat, sentimen, maupun pandangan. Salah satu metode yang dapat digunakan untuk menganalisis sentimen di platform YouTube terhadap pasangan calon presiden dan wakil presiden dalam

Pemilu Presiden dan Wakil Presiden Indonesia 2024 yaitu dengan menggunakan analisis sentimen berbasis *lexicon*.

Analisis sentimen berbasis *lexicon* atau berbasis kosa kata yaitu salah satu pendekatan di dalam analisis sentimen dengan memanfaatkan pembobotan kosa kata atau frasa. Pembobotan tersebut yaitu dengan melabeli kata atau frasa yang ada dengan positif atau negatif (Bonta dkk., 2019). Untuk meningkatkan hasil analisis sentimen berbasis *lexicon* dapat dikombinasikan dengan metode klasifikasi *Support Vector Machine* (SVM).

Terdapat beberapa penelitian sebelumnya yang melakukan analisis sentimen berbasis *lexicon* dengan menggunakan metode *Support Vector Machine* (SVM). Misalnya, penelitian berjudul “*Combination of Support Vector Machine and Lexicon-Based Algorithm in Twitter Sentiment Analysis*” yang ditulis oleh Muhammadi dkk. (2022). Penelitian ini berfokus pada penggunaan metode *lexicon* untuk melakukan pelabelan pada data dan kemudian hasil sentimen beserta dengan data yang telah terlabel digunakan oleh SVM. Hasil dari penelitian ini yaitu semua *kernel* yang diuji coba di dalam algoritma SVM berhasil mengungguli tingkat akurasi dari dua algoritma lain untuk klasifikasi seperti *Naïve Bayes* dan *K-Nearest Neighbors* (KNN). Selanjutnya, terdapat penelitian lain yang berjudul “*Sentiment Analysis for Online Learning using The Lexicon-Based Method and The Support Vector Machine Algorithm*” yang ditulis oleh Anam dkk. (2023). Hasil dari penelitian tersebut yaitu penggunaan algoritma SVM berbasis *lexicon* dapat meningkatkan akurasi sebesar 5% dibandingkan dengan tidak menggunakan basis *lexicon*. Disebutkan bahwa di dalam penelitian tersebut akurasi dari penggunaan algoritma SVM saja akan mencapai tingkat akurasi sebesar 90% dan apabila dikombinasikan dengan *lexicon* akan menambah tingkat akurasi tersebut hingga 95%. Selain itu, terdapat penelitian yang berjudul “*Sentiment Analysis for Predicting Customer Reviews using a Hybrid Approach*” yang ditulis oleh Rajeswari dkk. (2020). Fokus dari penelitian tersebut yaitu membahas perbandingan penggunaan metode analisis sentimen antara pendekatan tanpa metode *lexicon* dengan pendekatan *hybrid*. Pendekatan *hybrid* yang dimaksud di penelitian tersebut yaitu menggunakan gabungan antara metode *machine learning* seperti *Naïve Bayes*, SVM, *Logistic Regression*, dan *Decision*

Tree dengan metode *lexicon*. Hasil dari penelitian tersebut yaitu penggunaan metode *machine learning* yang digabungkan dengan metode *lexicon* akan menunjukkan peningkatan akurasi daripada penggunaan metode *machine learning* yang tidak menggunakan metode *lexicon*.

Berdasarkan penelitian yang telah dijelaskan sebelumnya dapat disimpulkan bahwa algoritma SVM memiliki performa yang baik, khususnya bila digabungkan dengan metode *lexicon*. Penggunaan metode *lexicon* untuk pelabelan data yang digabungkan dengan algoritma SVM juga memiliki peran dalam menciptakan akurasi sentimen yang lebih baik. Oleh karena itu, di dalam penelitian ini diterapkan analisis sentimen menggunakan algoritma SVM dengan metode *lexicon*. Penelitian ini mengklasifikasikan sentimen positif dan negatif dari komentar video YouTube pada debat kepresidenan untuk mengetahui sentimen masyarakat kepada masing-masing pasangan kandidat calon presiden dan wakil presiden.

Selain itu, untuk meningkatkan pemahaman dari analisis sentimen yang dilakukan, pada penelitian ini juga diimplementasikan *co-occurrence network* untuk dapat membantu mengidentifikasi hubungan antar kata yang sering muncul bersama di dalam sebuah komentar. Implementasi *co-occurrence network* membantu dalam proses visualisasi yang mudah dipahami, sehingga dapat diketahui topik-topik yang saling berkaitan di dalam komentar. Hal tersebut dapat memberikan pengetahuan lebih mengenai persepsi masyarakat terhadap para pasangan kandidat.

Penerapan analisis sentimen menggunakan metode SVM diharapkan menghasilkan berbagai sentimen masyarakat mengenai pasangan calon presiden dan wakil presiden yang bersaing di Pemilu Presiden dan Wakil Presiden Indonesia 2024. Berdasarkan hasil sentimen tersebut, penelitian ini diharapkan dapat digunakan untuk mengetahui bagaimana persepsi publik terhadap masing-masing pasangan kandidat calon presiden dan wakil presiden berdasarkan hasil debat.

I.2 Perumusan Masalah

Berdasarkan latar belakang, rumusan masalah dari penelitian ini yaitu:

- a. Bagaimana implementasi analisis sentimen berbasis *lexicon* dengan metode *Support Vector Machine* (SVM) pada platform YouTube mengenai debat kepresidenan di Pemilu Presiden dan Wakil Presiden Indonesia 2024?
- b. Bagaimana pengaruh penggunaan metode *lexicon* terhadap performa algoritma SVM pada model?
- c. Bagaimana komunitas kata yang terbentuk atas kata-kata yang muncul di dalam sentimen positif dan negatif?

I.3 Tujuan Penelitian

Berdasarkan rumusan masalah yang dijelaskan sebelumnya, penelitian ini bertujuan untuk:

- a. Mengetahui implementasi analisis sentimen berbasis *lexicon* dengan metode *Support Vector Machine* (SVM) pada platform YouTube pada debat kepresidenan Pemilu Presiden dan Wakil Presiden Indonesia 2024.
- b. Mengetahui pengaruh penggunaan metode *lexicon* pada algoritma SVM pada model analisis sentimen.
- c. Mengetahui komunitas kata yang terbentuk atas kata-kata yang muncul di dalam sentimen positif dan negatif.

I.4 Batasan Penelitian

Batasan dari penelitian ini yaitu:

- a. Data yang diambil dalam penelitian ini yaitu berasal dari komentar video debat presiden dan wakil presiden yang ditayangkan secara langsung dari kanal YouTube KPU RI dan Najwa Shihab.
- b. Komentar yang diambil dibatasi dari tanggal penayangan video debat, yaitu 12 Desember 2023, 22 Desember 2023, 7 Januari 2024, 21 Januari 2024, dan 4 Februari 2024, hingga berakhir pada tanggal 10 Februari 2024.

I.5 Manfaat Penelitian

Adapun manfaat penelitian ini yaitu:

1. Bagi peneliti, penelitian ini bermanfaat untuk dapat menerapkan pengetahuan yang telah dipelajari dan menambah pengalaman dalam melakukan pengelolaan data beserta dengan analisisnya.
2. Bagi keilmuan, penelitian ini bermanfaat untuk menambah literasi mengenai pendekatan *hybrid* dalam melakukan analisis sentimen, khususnya pada analisis sentimen berbasis *lexicon* dengan metode *Support Vector Machine* (SVM).

I.6 Sistematika Penulisan

Penelitian ini diuraikan dengan sistematika penulisan sebagai berikut:

Bab I Pendahuluan

Bab ini berisi uraian mengenai konteks permasalahan, latar belakang penelitian, perumusan masalah, tujuan penelitian, batasan penelitian, manfaat penelitian, dan sistematika penulisan.

Bab II Tinjauan Pustaka

Bab ini berisi penjelasan mengenai teori-teori yang relevan dengan permasalahan yang diteliti serta dibahas tentang hasil dari penelitian terdahulu yang berkaitan dengan topik penelitian.

Bab III Metodologi Penelitian

Bab ini berisi mengenai metodologi yang digunakan di dalam penelitian beserta dengan penjelasan mengenai sistematika penelitian yang dilakukan.

Bab IV Analisis dan Perancangan

Bab ini menjelaskan mengenai proses-proses yang dilakukan dari penelitian yang mencakup *initialization*, *data selection*, *data preprocessing*, *data transformation*, *data mining*, dan *interpretation/evaluation*.

Bab V Implementasi dan Pengujian

Bab ini menjelaskan mengenai hasil dari pembuatan model beserta dengan evaluasi performa model dari masing-masing pasangan kandidat.

Bab VI Kesimpulan dan Saran

Bab ini menjelaskan mengenai kesimpulan dari penelitian yang dilakukan beserta jawaban dari pertanyaan penelitian yang disajikan pada Bab Pendahuluan. Saran penelitian diberikan pada bab ini untuk penelitian selanjutnya.