

BAB I

PENDAHULUAN

1.1 Latar Belakang

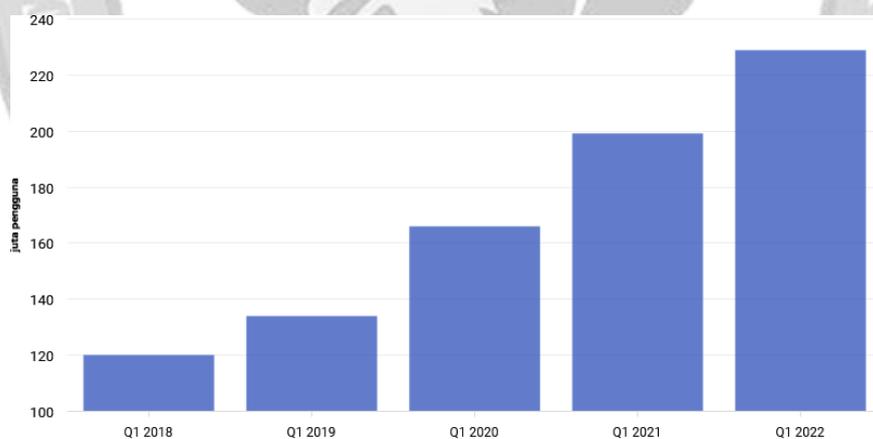
Media sosial (Medsos) adalah sebuah platform digital yang dibangun untuk melakukan interaksi sosial tanpa adanya batasan jarak dan waktu. Medsos ini dimungkinkan melakukan interaksi satu sama lain dimanapun dan kapanpun, tidak peduli jarak atau waktu yang ditempuh. Menurut KBBI medsos adalah laman atau aplikasi yang memungkinkan pengguna dapat membuat dan berbagi isi atau terlibat dalam jaringan sosial [1].

Twitter didirikan pada bulan Maret 2006 oleh Jack Dorsey, dan situs jejaring sosialnya diluncurkan pada bulan Juli. Sejak diluncurkan, *Twitter* telah menjadi salah satu dari sepuluh situs yang paling sering dikunjungi di internet, dan dijuluki dengan pesan singkat dari internet. Beberapa istilah dalam *Twitter* antara lain: *Mention*, adalah menyebut atau memanggil pengguna *Twitter* lain dalam sebuah tweet. *Mention* dilakukan dengan menuliskan '@' diikuti dengan nama pengguna lain. Hashtag, digunakan untuk menandai sebuah topik pembicaraan di *Twitter*. Penulisan hashtag dimulai dengan tanda '#' diikuti dengan topik yang sedang dibahas. Hashtag biasa digunakan untuk meningkatkan visibilitas tweet pengguna. Hashtag adalah cara untuk menandai sebuah topik pembicaraan di *Twitter*, maka trending topics adalah kumpulan dari topik pembicaraan yang sangat populer di *Twitter* [2].

Twitter merupakan jejaring sosial yang memungkinkan penggunanya untuk membaca dan mengirim pesan berbasis teks hingga 280 karakter. Tingginya popularitas *Twitter* menyebabkan layanan ini dimanfaatkan untuk keperluan dalam

berbagai aspek. Seperti ekonomi, politik, sosial dan budaya. Dalam sebuah politik kunci keberhasilan adalah mengumpulkan suara terbanyak. Sosial media khususnya banyak digunakan untuk kampanye pemilu. Selain itu juga *Twitter* digunakan oleh para kritikus untuk melakukan promosi dan berinteraksi dengan warganet.

Berdasarkan data dari sebuah website [3] Pengguna *Twitter* di kuartal I-2018 mencapai 120 juta pengguna, pada Gambar 1.1 dibawah. Kemudian jumlahnya naik 134 juta pengguna pada kuartal ke I-2019, pada kuartal ke I-2020 pengguna *Twitter* naik menjadi 166 juta orang. Selain itu kenaikan di kuartal I-2021 tercatat 199 juta pengguna. Dalam data terbaru jumlah pengguna *Twitter* mencapai 229 juta orang. Angka tersebut naik sebanyak 15.07% dibandingkan periode yang sama pada tahun sebelumnya (*Year-On-Year/YOY*). Tercatat pengguna terbanyak didunia berasal di Amerika Serikat (AS).



Gambar 1. 1 Pengguna *Twitter* di Q1 2018 - Q1 2022

Pemilu atau pemilihan umum yaitu proses memilih orang untuk dijadikan pengisi jabatan-jabatan politik tertentu, mulai dari Presiden, DPR, MPR dan lain lain. Pemilu menurut Undang-Undang Nomor 8 Tahun 2012 pasal 1 ayat 6 adalah sarana pelaksanaan kedaulatan rakyat dalam Negara Kesatuan Republik Indonesia

yang berdasarkan Pancasila dan Undang-Undang Dasar Negara Republik Indonesia tahun 1945 [4]. Pemilihan orang dilakukan secara terbuka, bebas namun rahasia. Untuk itu para calon harus memperkenalkan diri ke publik. Cara termurah dan mudah melakukannya tentu menggunakan medsos.

Dengan perkembangan teknologi yang terus berinovasi dan didukung oleh kemajuan internet melahirkan banyak medsos. Sebagai contoh lahirnya *Facebook*, *Instagram*, *Twitter*, *Tiktok* dan lain – lain. Pada umumnya medsos digunakan untuk menambah informasi, menambah wawasan tanpa adanya batasan ruang dan waktu. Akhir – akhir ini, medsos banyak digunakan dalam kampanye pemilu dan marketing bisnis. Selain sebagai media promosi beberapa juga sebagai media kritikan untuk lawan politik. Opini yang dilontarkan dari pengguna sosial media khususnya *Twitter* memicu adanya sebuah opini. Dalam opini tersebut ada sebuah sentiment.

Dalam aplikasi *Twitter* pengguna dapat leluasa membuat *tweet* atau membuat status sesuai dengan ekspresi yang sedang dirasakan. Selain itu juga dalam *Twitter*, pengguna dapat saling memberikan komentar dan suka terhadap status atau *tweet* orang lain. Sebagai contoh sebuah *tweet* “Sepertinya pak X Lebih cocok deh pasangannya dengan pak Y Sebagai calon presiden selanjutnya”. Status atau *tweet* dari pengguna tersebut merupakan sebuah opini. Dalam opini tersebut muncul sebuah sentimen terhadap calon calon presiden tertentu. Opini tersebut bisa berupa sebuah *tweet* atau status bisa juga berupa sebuah komentar. Sentiment tersebut nantinya akan dilakukan penilaian negatif, positif, atau netral.

Penilaian positif, negative dan netral dalam suatu opini dapat dilakukan secara manual. Dengan perkembangan jaman, penilaian opini dapat dilakukan

dengan menggunakan pembelajaran mesin atau *Machine Learning* selanjutnya dilakukan pengklasifikasian opini dari sebuah sumber data. Untuk melakukan pengklasifikasian digunakan fungsi dari *text mining* atau penambangan text. Dalam hal ini pengklasifikasian dalam bentuk dokumen.

Penggunaan medsos itu harus direncanakan, di komunikasikan dan di program untuk meningkatkan kredibilitas orang atau partai. Medsos telah memainkan peran penting dalam kampanye politik 2014, melalui platform seperti Facebook, *Twitter*, dan Youtube. Kandidat politik akan terus berinteraksi dengan pendukung dan menerima dukungan dalam bentuk sumbangan dan relawan. [5]. Berdasarkan penelitian Rehan dkk [6] dalam hasil temuan mereka disimpulkan bahwa pemanfaatan medsos dalam kampanye politik oleh kandidat calon kepala daerah Kabupaten Pesisir Selatan belum maksimal untuk mempengaruhi politik. Namun Amanda dkk [7] menyimpulkan bahwa medsos memberikan dampak yang signifikan.

Penelitian lain yang dilakukan oleh Safitri [8] Tentang “Analisis Sentimen Persepsi Masyarakat Terhadap Pemilu 2019 Pada Media Sosial *Twitter* Menggunakan Naive Bayes” dalam kesimpulannya terbentuk pola persepsi maka dilakukan proses modeling dengan klasifikasi Naive Bayes dan menemukan bahwa data set tweet pada penelitian memiliki pola persepsi negatif atau sentimen negatif sebesar 52% jauh lebih besar dari persepsi positif yaitu 18% dan Persepsi positif juga lebih kecil dari persepsi netral dimana persepsi netral memiliki nilai 31%. Tingkat akurasi klasifikasi Naive Bayes terhadap dataset training bernilai 81% dan 76% pada dataset testing, sedangkan rata-rata nilai precision untuk sentimen positif 86,65%, sentimen negatif 77,15%, dan sentimen netral bernilai 80,95% sedangkan

rata-rata nilai recall pada sentimen positif 36,8%, sentimen negatif 93,2% dan sentimen netral 86,8%. Penelitian ini dilakukan pada pemilu tahun 2019 data yang digunakan menggunakan dataset dari *Twitter* pada tahun 2018 – 2019.

Penelitian yang telah dilakukan oleh Amirul dkk [9] tentang kinerja KPU dipemilu 2019 dengan menggunakan 2 skenario. Skenario 1 menggunakan algoritma *K-Means* tanpa bantuan algoritma normalisasi kata dan algoritma *Stemming* didapatkan nilai akurasi 84%. Sedangkan dengan menggunakan scenario 2 menggunakan algoritma *K-Means*, algoritma *Levensthein Distance* sebagai normalisasi kata serta algoritma *Confix Stripping Stemmer* sebagai algoritma stemming mengalami peningkatan 86%. Dalam penelitiannya masyarakat pada umumnya mereka menggambarkan pendapat mereka menggunakan kata – kata singkatan yang tidak ada di KBBI, penggunaan fitur normalisasi dan stemming peneliti ini dapat memberikan informasi lebih jelas dan lebih akurat. Dengan tingkat akurasi 84% tanpa normalisasi dan 86% dengan menggunakan normalisasi.

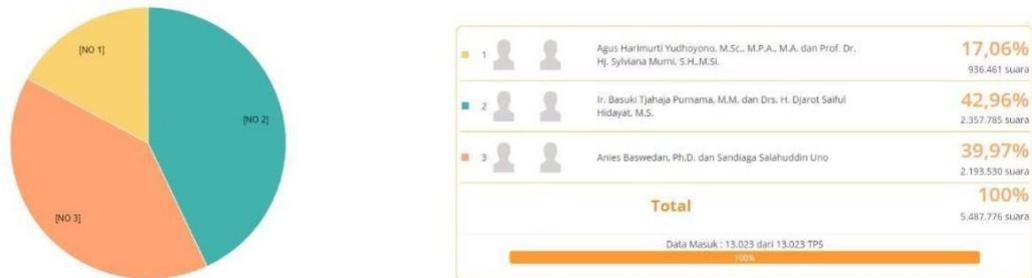
Berdasarkan kedua penelitian yang dilakukan oleh Safitri [8] dan Amirul dkk [9] terhadap analisis sentiment persepsi masyarakat terhadap pemilu 2019 pada media sosial *Twitter*, dapat dilihat bahwa masyarakat pada umumnya memiliki persepsi negatif terhadap pemilu tersebut. Pada penelitian Safitri, terdapat pola persepsi negatif sebesar 52% yang jauh lebih besar dibandingkan persepsi positif yaitu 18% dan persepsi netral 31%. Sedangkan pada penelitian Amirul dkk, tingkat akurasi dari algoritma *K-Means* sebesar 84% tanpa normalisasi dan 86% dengan menggunakan normalisasi.

Namun, perlu diingat bahwa analisis sentiment ini hanya berdasarkan data tweet yang diambil dari tahun 2018-2019 dan tidak dapat dikatakan berkorelasi

langsung dengan proses pemilu 2019. Faktor-faktor lain seperti tingkat partisipasi masyarakat, tingkat kecukupan informasi, dan tingkat kejujuran dalam pemungutan suara juga harus dipertimbangkan dalam membicarakan kevalidan proses pemilu. Oleh karena itu, untuk memastikan apakah proses pemilu 2019 sesuai dengan analisis sentiment yang dilakukan, perlu dilakukan penelitian lebih lanjut yang mengkaji secara komprehensif dan mendalam.

Penelitian Ghulam [10] tentang analisis sentimen calon gubernur DKI Jakarta pada tahun 2017. Dalam penelitian ini data set yang digunakan sama yaitu cuitan status dalam *Twitter*. Dalam penelitian yang dilakukan oleh Ghulam data tweet yang digunakan yaitu 100 tweet dengan kata kunci AHY, 100 tweet dengan kata kunci Ahok, dan 100 tweet dengan kata kunci Anies. Dalam penelitian tersebut penelitian sebelumnya mengambil data tweet kemudian dilakukan penilaian sentiment dengan menggunakan algoritma *Naïve Bayes Classifier (NBC)* untuk klasifikasi AHY, dengan nilai rata-rata akurasi mencapai 95% nilai presisi 95% nilai *recall* 95% nilai *TP rate* 96,8% dan nilai *TN rate* 84,6% dan menggunakan algoritma *Support Vector Machine (SVM)* dengan klasifikasi AHY nilai rata-rata akurasi yaitu 90% nilai presisi 89,9% nilai *recall* 90% nilai *TP Rate* 98,4% nilai *TN rate* 38.5% dalam penelitian tersebut metode klasifikasi *Naïve Bayes Classifier (NBC)* lebih tinggi akurasinya dibandingkan dengan *Support Vector Machine (SVM)*.

Hasil Hitung TPS (Form C1) Provinsi Dki Jakarta



Gambar 1. 2 Hasil Hitung TPS (Form C1) Provinsi Dki Jakarta

Pada gambar 1.2, hasil perhitungan TPS (*From C1*) Provinsi Jakarta dengan Nomor urut 1. Yaitu Agus Harimurti Yudhoyono, M.Sc., M.P.A., M.A. dan Prof. Dr. Hj. Sylviana Murni, S.H., M.Si. mendapatkan poling sebesar 17,06% dengan hasil suara 936.461 suara. Pada nomor urut dua yaitu Ir. Basuki Tjahaja Purnama, M.M. dan Drs. H. Djarot Saiful Hidayat, M.S. mendapatkan poling sebesar 42,92% dengan hasil suara 2.357.785 suara. Dan pada nomor urut ketiga yaitu Anies Baswedan, Ph.D. dan Sandiaga Salahuddin Uno mendapatkan poling sebesar 39,97% dengan hasil suara 2.193.530, Total suara yang didapat yaitu 5.487.776 Suara. Berdasarkan hasil pemilu putaran pertama tersebut pemenang pemilihan Gubernur 2017 putra pertama dimenangkan oleh Ir. Basuki Tjahaja Purnama, M.M. dan Drs. H. Djarot Saiful Hidayat, M.S. Penelitian yang dilakukan oleh Ghulam sesuai dengan hasil aktual pemilu gubernur DKI Jakarta pada tahun 2017 dimenangkan oleh Basuki atau ahok. Hasil penelitian menunjukkan AHY mendapatkan score yang tinggi, hal tersebut dapat dilihat bahwa metode klasifikasi Naive Bayes Classifier yang digunakan dalam penelitian memiliki tingkat akurasi yang tinggi, dengan nilai rata-rata akurasi sebesar 95%. Ini menunjukkan bahwa algoritma tersebut mampu memberikan hasil yang baik dalam mengklasifikasi

sentimen pada data tweet. Meskipun pada hasil *real* pemenang pemilu tersebut adalah Ir. Basuki Tjahaja Purnama, M.M. dan Drs. H. Djarot Saiful Hidayat, M.S. Penelitian D. O. Manalu, M. C. L. Tobing and M. Yohanna [11] terhadap wacana penunandaan pemilu dengan metode *Support Vector Machine* dengan sampel yang diambil sebanyak 100 *Tweet*, terbukti dengan metode SVM dapat menghasilkan klasifikasi pada tingkat akurasi sebesar 90%. Dengan hasil pelabelan sentiment dengan manual sebanyak 100 sampel data *Tweet*, diperoleh sentiment positif 40 data *Tweet* dan sentiment negatif sebanyak 60 data *Tweet*. Dalam penelitian tersebut respon dari masyarakat cenderung negatif dan kurang sependapat dengan wacana penunandaan pelaksanaan pemilu pada tahun 2024.

Berdasarkan penelitian yang dilakukan oleh Raihan dkk [12] penelitian dapat dilihat bahwa warganet memiliki sentimen positif dari dataset yang digunakan. Hasil penelitian tersebut memiliki akurasi 62,5% dan AUC 0,65 dengan dataset Ridwan. Kemudian untuk dataset Prabowo dengan akurasi 60% dan AUC 0,92. Selanjutnya dari dataset Anies memiliki akurasi 71,43% dan AUC 1,0. Pada dataset yang terakhir yaitu Ganjar dengan akurasi sebesar 73,68% dan AUC 0,74. Hasil dari nilai sentimen yang dihasilkan melalui algoritme naïve bayes juga beragam dengan nilai terbesar dari dataset Ridwan dengan nilainya 1, kemudian dataset Ganjar 0,9365, dataset Anies senilai 0,931, dan dataset Prabowo 0,3617. Selain dari hasil tersebut digunakan juga nilai dari skor sentimen dengan nilai terbesar dari dataset Ridwan dengan nilainya adalah 4,878, kemudian dataset Anies dengan nilai 4, dataset Ganjar dengan nilai 3,879, dan dataset Prabowo dengan nilai skor sebesar 2,533.

Berdasarkan hasil penelitian D.O. Manalu, M.C.L. Tobing dan M. Yohanna [11], masyarakat cenderung memiliki sentiment negatif terhadap wacana penundaan pemilu tahun 2024. Ini dibuktikan dengan hasil pelabelan sentiment manual pada 100 sampel data Tweet, diperoleh 40 data Tweet sentiment positif dan 60 data Tweet sentiment negatif. Berdasarkan penelitian Raihan dkk [6], hasil analisis sentimen pada dataset Ridwan memiliki akurasi 62,5% dan nilai sentimen yang paling besar, yaitu 1. Sementara untuk dataset Prabowo memiliki akurasi 60% dan nilai sentimen terendah, yaitu 0,3617. Dataset Anies memiliki akurasi 71,43% dan nilai sentimen sebesar 0,931, dan dataset Ganjar memiliki akurasi 73,68% dan nilai sentimen sebesar 0,9365. Dengan demikian, hipotesis dari hasil penelitian tersebut adalah bahwa sentimen warganet terhadap wacana penundaan pemilu tahun 2024 bervariasi, namun cenderung negatif. Hal tersebut menjadi acuan untuk penelitian yang akan penulis kembangkan.

Tahun 2024 akan diadakannya pemilu di Indonesia. Oleh karena itu dapat diprediksi bahwa tahun 2023 menjadi tahun komunikasi politik di Indonesia dan suhu politik akan memanas. Medsos merupakan media komunikasi yang akan hangat dan sangat mungkin dipakai untuk mengkomunikasikan politik untuk membentuk opini publik seperti. Rasa ingin tahu penulis tentang kandidat calon presiden yang akan muncul pada tahun 2024. Maka penulis tertarik untuk melakukan penelitian “Analisis Sentimen Terhadap Calon Presiden 2024” pada sosial media twitter Menggunakan Algoritma Support Vector Machine.

1.2 Identifikasi Permasalahan

Dari beberapa uraian yang disampaikan di latar belakang di atas, maka dapat diidentifikasi masalah-masalah sebagai berikut:

- Banyak ragam media sosial yang bisa dipakai kelompok kampanye untuk menyebarkan berita.
- Munculnya hoaks atau berita palsu yang dapat menyesatkan masyarakat dan mempengaruhi pemilihan.
- Sosial media dapat dipakai menyebar luaskan pandangan yang sangat berbeda sehingga membuat masyarakat terpecah belah dan tidak bisa bekerja sama dalam menyelesaikan masalah. Sehingga dapat memunculkan polarisasi ekstrim di masyarakat.
- Kerahasiaan data pribadi yang tidak terjamin sehingga memungkinkan beberapa kandidat atau kelompok kampanye dapat mengumpulkan data pribadi pengguna sosial media untuk digunakan dalam kampanye, tanpa persetujuan atau izin dari pengguna.
- Cuitan berupa *Tweet* negative yang menyebar provokasi dapat menyebabkan kerusuhan yang merugikan semua pihak.
- Menurunnya tingkat kepercayaan masyarakat terhadap beberapa tweet negatif yang tidak dapat dibuktikan dapat menurunkan tingkat kepercayaan masyarakat terhadap sosial media, sehingga mempengaruhi tingkat partisipasi masyarakat dalam pemilu.

1.3 Batasan Masalah

Batasan masalah pada penelitian ini dengan hanya mengambil tweet dalam Bahasa Indonesia dari satu media sosial yaitu Twitter dengan topik calon presiden 2024 Republik Indonesia. Data dianalisis dengan model pembelajaran mesin atau *machine learning* dengan menggunakan *tools* seperti *Python*. Rentang waktu pengambilan data *tweet* yaitu pada 01 Januari 2022 hingga 01 Januari 2023 Dengan dataset berjumlah 16850. Selain itu juga penulis menggunakan kata kunci “Pemilu 2024” untuk data *tweet* yang digunakan untuk sekitar DKI Jakarta.

1.4 Rumusan Masalah

Berdasarkan latar belakang yang telah di kemukakan, maka dapat merumuskan permasalahan sebagai berikut :

1. menganalisis sentimen opini publik terhadap pemilu 2024 dalam pemilihan presiden di tahun 2024 melalui sosial media twitter ?
2. Bagaimana perbandingan dari akurasi model *Naïve Bayes Classifier* dengan model *Support Vector Machine*?

1.5 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah dirumuskan di atas, tujuan penelitian sebagai berikut :

1. Mengetahui sentiment opini public terhadap calon presiden dalam pemilihan presiden di tahun 2024 melalui social media twitter dengan metode *Text Mining*.
2. Mengetahui tingkat keakuratan model *Naïve Bayes Classifier* dan *Support Vector Machine* terhadap calon presiden 2024

1.6 Sistematika Penulisan

Penulisan dalam penelitian ini terbagi menjadi sedikitnya dalam lima bab.

Adapun sistematikanya sebagai berikut:

1. Bab I

Membahas dasar latar belakang Skripsi, permasalahan yang diajukan, tujuan dari Skripsi, keterbatasan dalam pelaksanaan Skripsi, dan kegunaan Skripsi.

Bab ini juga menguraikan struktur penulisan Skripsi yang digunakan.

2. Bab II

menguraikan dasar teoretik yang diterapkan dalam Skripsi ini. Beberapa teori yang dibahas meliputi Pemilu 2024, analisis sentimen, pandangan, Twitter, penambahan teks, representasi dokumen, pendekatan analisis sentimen, model SVM, Model Naive Bayes, pemodelan tema, dan penilaian model. Selain itu, Bab II juga mencantumkan literatur yang menjadi acuan dalam penulisan tesis ini.

3. Bab III

Bab III menguraikan metode yang diterapkan untuk mencapai tujuan Tesis. Metode ini mencakup langkah-langkah dalam Skripsi ini, mulai dari proses pengumpulan data sampai analisis sentimen dan pemodelan tema.

4. Bab IV

Bab IV membahas tentang temuan dari analisis sentimen dengan menggunakan model SVM serta pemodelan tema menggunakan LDA.

Selanjutnya, dibahas pula hasil analisis terkait pandangan masyarakat

mengenai Analisis Sentimen Pemilu 2024 dan hubungannya dengan berita yang beredar di masyarakat.

5. Bab V

Kesimpulan dan Saran. Pada bab ini menjelaskan tentang kesimpulan berdasarkan hasil dari penelitian yang diperoleh, serta saran-saran konstruktif yang perlu dikembangkan untuk penelitian berikutnya sehingga penelitian berikutnya menjadi lebih baik.

