

BAB I

PENDAHULUAN

1.1 Latar Belakang

Internet adalah salah satu sarana untuk mengakses informasi dan komunikasi yang dapat diakses oleh seseorang dimana saja dan kapan saja. Seiring dengan berjalannya kemajuan teknologi, penggunaan internet sebagai sarana komunikasi juga bertumbuh sangat pesat. Indonesia adalah salah satu negara dengan pengguna internet terbesar di dunia. Berdasarkan hasil survey dari APJII (Asosiasi Penyelenggara Jasa Internet Indonesia) pada tahun 2022, jumlah penduduk yang terkoneksi internet mencapai 210.026.769 jiwa dengan peningkatan persentase penetrasi internet sebesar 12,22% dari tahun 2018 (APJII, 2022). Faktor meningkatnya pengguna internet di Indonesia disebabkan oleh beragam alasan, berdasarkan hasil survey dari APJII (Asosiasi Penyelenggara Jasa Internet Indonesia) mengakses media sosial (termasuk mengakses Facebook/Whatapp/Telegram/Twitter/Instagram/Youtube/dll) adalah alasan yang paling banyak digunakan masyarakat untuk mengakses internet dengan persentase skala penilaian mencapai 98,02%. Dengan berkembangnya pengguna internet di Indonesia menunjukkan bahwa penggunaan media sosial akan terus berkembang seiring dengan berkembangnya teknologi (APJII, 2022).

Penggunaan internet semakin meluas dengan banyaknya *platform* yang menyediakan berbagai layanan kepada masyarakat untuk mengakses internet. Media sosial menjadi salah satu platform yang banyak digunakan untuk berkomunikasi, dan mengakses informasi dengan cepat. Informasi yang terdapat pada media sosial sangat beragam, seperti komentar, kritikan, hingga opini. Twitter menjadi salah satu platform yang banyak digunakan oleh masyarakat Indonesia. Dilansir dari *We Are Social*, pengguna Twitter di Indonesia mencapai 18,45 juta pada tahun 2022 (We Are Social, 2022). Hal ini membuktikan bahwa terdapat banyak sekali sentimen masyarakat Indonesia yang terdapat pada *platform* Twitter yang dapat dijadikan sebagai studi kasus.

Indonesia merupakan negara demokrasi yang mana peran rakyat sangat menentukan dalam memilih pemimpin negara yang dilaksanakan melalui pemilihan umum atau yang sering disebut pemilu. Pemilihan Presiden atau pilpres merupakan pesta rakyat yang selalu menjadi perbincangan hangat ketika menjelang pemilu. Banyak sekali opini pro dan kontra dari berbagai kalangan masyarakat mengenai calon presiden dan wakil presiden semasa menjelang pemilu. Tak luput juga pilpres yang akan dilaksanakan pada tahun 2024 mendatang. Dengan semakin dekatnya waktu pemilu, semakin jelas juga calon calon presiden dan wakil presiden yang akan melaksanakan kampanye.

Beberapa tokoh publik sudah mulai banyak diprediksi sebagai calon presiden dalam beberapa sentimen pada *platform* Twitter seperti Ganjar Pranowo, Prabowo Subianto, dan Anies Baswedan. Tokoh publik lainnya yang ikut diprediksi sebagai calon wakil presiden dalam *Twitter* yaitu Erick Thohir, Agus Harimurti Yudhoyono (AHY), dan Airlangga Hartanto. Nama-nama diatas akan dijadikan sebagai *keyword* dalam klasifikasi sentimen pengguna *Twitter* mengenai pilpres 2024 agar terlihat citra para tokoh publik tersebut apakah positif atau negatif di mata masyarakat berdasarkan sentimen masyarakat di Twitter.

Analisis sentimen sangat diperlukan dalam memantau media sosial. Terlebih lagi dalam memantau sentimen masyarakat mengenai pilpres 2024. Di media sosial, analisis sentimen berguna untuk mendapatkan gambaran umum dari suatu opini. Pemilihan presiden 2024 akan dilaksanakan di tahun mendatang, namun sudah banyak sekali bermacam-macam sentimen dari masyarakat baik itu sentimen positif maupun negatif untuk para calon presiden Indonesia. Namun bagi masyarakat yang awam terhadap politik maupun yang tidak mengikuti perkembangan di media sosial, mereka masih belum menemukan pilihan presiden yang tepat di hati untuk memimpin negara Indonesia.

Berdasarkan survei capres 2024 yang dilakukan oleh Lembaga Survei Nasional (LSN) pada tanggal 29 Oktober hingga 2 November 2022, didapatkan bahwa dari 1.230 responden 26,5% diantaranya masih tidak tahu atau tidak menjawab siapa calon presiden Indonesia 2024. Hasil survei terbaru lainnya yang dilakukan oleh Lembaga Survei Indonesia (LSI) pada tanggal 7-11 Januari 2023, dari 1.221 responden 17% diantaranya masih tidak tahu atau tidak menjawab siapa calon

presiden Indonesia 2024 nantinya. Maka dari itu analisis sentimen diperlukan untuk menyediakan informasi bagi masyarakat mengenai mayoritas pilihan masyarakat terhadap calon presiden berdasarkan opini publik dari Twitter.

Sebelumnya, telah dilakukan penelitian mengenai analisis sentimen terhadap calon presiden Indonesia dengan berbagai metode. Diantaranya yaitu penggunaan algoritma *Support Vector Machine* (SVM) dan *K-Nearest Neighbor* (KNN) dengan *word embedding* TF-IDF terhadap calon presiden 2019, dengan hasil akurasi pada SVM sebesar 76,5% dan hasil akurasi pada KNN sebesar 68,3% (Firmansyah, 2019). Penelitian lainnya yang menerapkan metode Naïve Bayes terhadap bakal calon presiden 2024, dengan hasil akurasi tertinggi sebesar 73,68% untuk *dataset* Ganjar Pranowo hingga akurasi terendah sebesar 60% untuk *dataset* Prabowo Subianto (Fais Sya' bani et al., 2022).

Sebelumnya juga telah dilakukan penelitian yang membahas analisis sentimen menggunakan metode *Long Short-Term Memory* (LSTM). Seperti pada penelitian analisis sentimen *tweet* mengenai COVID-19 menggunakan *Word Embedding* dan metode *Long Short-Term Memory* (LSTM) dengan hasil perbandingan dari metode LSTM, RNN dan Naïve Bayes. Didapatkan hasil akurasi metode LSTM sebesar 81%, akurasi metode RNN sebesar 74% dan akurasi metode Naïve Bayes sebesar 71% (Rahman et al., 2021). Penelitian lainnya yang terkait yaitu analisis sentimen publik twitter terhadap pemilu 2024 menggunakan LSTM, pelabelan dengan *lexicon based feature* dan tanpa melakukan *hyperparameter tuning*, didapatkan hasil akurasi sebesar 78% (Firdlous & Andrian, 2023).

Melihat dari hasil penelitian terdahulu, algoritma LSTM memiliki hasil akurasi paling tinggi dibandingkan dengan algoritma lainnya. Ini dikarenakan LSTM memiliki kelebihan dalam melakukan analisis sentimen, yaitu mampu mengingat informasi dari data yang telah diproses dan menggunakan kemampuan memori tersebut untuk memahami data input lebih baik (Sosa, 2017). Kemampuan memori ini hanya dimiliki oleh algoritma LSTM yang merupakan pengembangan dari algoritma *Recurrent Neural Network* (RNN) dalam mengatasi masalah *long-term dependency* dalam mengingat data jangka panjang.

Berdasarkan pada latar belakang dan hasil dari penelitian diatas penulis akan melakukan penelitian Analisis Sentimen Pengguna Twitter Mengenai Calon

Presiden Indonesia Tahun 2024 Menggunakan Algoritma LSTM. Penulis juga akan melakukan *data scraping* Twitter dengan menggunakan *library* SNScrape berbasis Python yang akan mengumpulkan seluruh data *tweet* dari seluruh pengguna Twitter yang nantinya akan dikumpulkan kedalam suatu *dataframe* agar data bisa diolah. Diharapkan dengan adanya penelitian ini dapat menghasilkan hasil akurasi yang lebih baik dengan mempertimbangkan beberapa atribut seperti *word embedding*, pelabelan teks hingga *hyperparameter tuning*.

1.2 Identifikasi Masalah

Bedasarkan paparan latar belakang diatas, maka dapat diidentifikasi masalah dalam penelitian ini antara lain sebagai berikut:

1. Banyaknya pengguna Twitter di Indonesia berdampak kepada polaritas sentimen pengguna Twitter khususnya mengenai calon presiden 2024.
2. Masih banyak masyarakat yang kesulitan untuk memutuskan pilihan calon presiden Indonesia 2024.
3. Model LSTM pada penelitian terdahulu tidak mempertimbangkan *hyperparameter tuning* sehingga membuat hasil akurasi model kurang optimal.

1.3 Ruang Lingkup

Ruang lingkup pada penelitian ini adalah sebagai berikut:

1. Penelitian ini akan dilakukan dengan menggunakan data *tweet* yang diperoleh dari pengguna Twitter dengan melakukan *data scraping* menggunakan SNScrape.
2. Model yang akan dibuat adalah pemodelan klasifikasi sentimen menggunakan algoritma LSTM.
3. Data yang akan digunakan pada penelitian ini adalah data *tweet* dari pengguna Twitter mengenai pilpres 2024 dan tokoh publik yang terlibat.

1.4 Rumusan Masalah

Berdasarkan identifikasi masalah yang telah dilakukan, maka dirumuskan masalah yang akan dibahas dalam penelitian ini sebagai berikut:

1. Bagaimana membangun model dan menentukan *hyperparameter tuning* yang tepat untuk menghasilkan akurasi yang baik pada model LSTM dalam menganalisis sentimen dari *tweet* mengenai calon presiden Indonesia tahun 2024?
2. Bagaimana mengevaluasi model LSTM dalam menganalisis sentimen dari *tweet* mengenai calon presiden Indonesia tahun 2024?

1.5 Batasan Masalah

Untuk memberikan ruang lingkup yang jelas dalam pembahasan, maka perlu dilakukan pembatasan masalah dalam penelitian ini sebagai berikut:

1. Data *tweet* yang digunakan hanya berasal dari pengguna Twitter yang aktif mulai dari 1 Februari 2022 hingga 28 Februari 2023.
2. Analisis sentimen hanya dilakukan terhadap *tweet* yang berisi opini atau pendapat mengenai calon presiden Indonesia 2024.
3. Model algoritma LSTM yang hanya diterapkan pada *tweet* bahasa Indonesia.
4. Klasifikasi sentimen yang dilakukan hanya klasifikasi sentimen positif dan negatif.
5. Hasil analisis sentimen tidak mempertimbangkan konteks atau keterkaitan *tweet* dengan *tweet* lain, hanya melihat sentimen dari *tweet* tersebut secara individual.

1.6 Tujuan Penelitian

Berdasarkan masalah yang telah diuraikan diatas maka tujuan penelitian yang akan dicapai adalah sebagai berikut:

1. Untuk membangun model dan menentukan *hyperparameter tuning* yang tepat agar menghasilkan akurasi yang baik pada model LSTM dalam

menganalisis sentimen dari *tweet* mengenai calon presiden Indonesia tahun 2024?

2. Untuk mengevaluasi model LSTM dalam menganalisis sentimen dari *tweet* mengenai calon presiden Indonesia 2024.

1.7 Manfaat Penelitian

Dengan diterapkannya tujuan penelitian diatas, maka diharapkan manfaat penelitian yang diperoleh setelah penelitian ini adalah:

1. Memberikan gambaran opini publik terkait calon presiden Indonesia 2024 yang akan dipilih berdasarkan sentimen positif dan negatif dari pengguna Twitter.
2. Membantu masyarakat dalam menentukan pilihan presiden berdasarkan hasil klasifikasi sentimen mengenai calon presiden Indonesia 2024.
3. Menjadi referensi tambahan dalam menganalisis pilihan calon presiden Indonesia 2024 disamping mengandalkan elektabilitas.