

BAB I.

LATAR BELAKANG

Indonesia adalah negara demokrasi yang ditunjukkan oleh sistem pemilihan umum yang memilih presiden setiap lima tahun sekali. Seluruh masyarakat Indonesia sangat menantikan pemilihan presiden, yang merupakan salah satu peristiwa politik yang paling penting. Dalam kebanyakan kasus, calon presiden, juga dikenal sebagai Capres, dipilih oleh partai politik atau gabungan dari beberapa partai politik. Keputusan ini didasarkan pada seberapa populer mereka di mata masyarakat. Twitter adalah salah satu jejaring sosial yang paling sering digunakan oleh masyarakat untuk menyampaikan pendapatnya terhadap Capres. Penelitian tentang analisis sentimen Capres 01 di media sosial tidak dapat diabaikan karena tidak ada pemahaman yang mendalam tentang persepsi dan opini masyarakat yang beredar secara online. Ini dapat berdampak pada efektivitas strategi kampanye dan membantu mereka menilai reputasi diri dan menanggapi masalah baru.

Twitter dan jenis jejaring sosial lainnya telah berkembang menjadi ruang publik virtual yang penuh dengan percakapan dan komunikasi informasi. Dengan banyaknya pengguna jejaring sosial Twitter, status seseorang dapat diubah menjadi data sentimen untuk diproses dan dianalisis oleh para peneliti. Analisis sentimen dari pengolahan status atau tweet pengguna Twitter menunjukkan komentar masyarakat terhadap calon presiden dan sentimen dari tweet tersebut. Tujuan analisis sentimen adalah untuk menganalisis, mengidentifikasi, dan mengungkapkan pendapat atau sentimen dalam teks. Analisis sentimen juga merupakan proses yang berfungsi untuk mengidentifikasi pendapat atau sentimen dari isi dataset, seperti teks tentang topik atau peristiwa yang positif, negatif, atau netral (Aftab dkk., 2023).

Ada banyak cara untuk menganalisis sentimen, yang umum digunakan adalah dengan menggunakan *machine learning* dan *Deep Learning*. *machine learning* memiliki metode yang cukup umum digunakan yaitu Naïve Bayes dan Support Vector Machines (SVM), sedangkan *Deep Learning* memiliki model yang cukup

umum digunakan untuk mengolah data berbentuk teks yaitu CNN. Pada penelitian yang dilakukan oleh Fais Sya' bani dkk. (2022) telah dilakukan analisis sentimen menggunakan metode Naïve Bayes. Penelitian ini mengevaluasi algoritma Naïve Bayes dengan akurasi 73,68%. Keunggulan dari penelitian ini adalah evaluasi akurasi dan AUC (Area Under Curve) yang dilakukan pada setiap dataset memberikan gambaran tentang kinerja algoritma Naïve Bayes dalam menganalisis sentimen. Riset ini menunjukkan bahwa algoritma Naïve Bayes menghasilkan tingkat akurasi yang cukup rendah. Rahman dkk. (2021) melakukan Penelitian yang juga menganalisis sentimen menggunakan algoritma LSTM menghasilkan evaluasi akurasi sebesar 81%, keunggulan penelitian ini adalah memiliki akurasi yang lebih baik dibandingkan algoritma Naïve Bayes. Penelitian kali ini akan menggunakan algoritma CNN untuk mencapai nilai akurasi yang lebih tinggi dari penelitian sebelumnya. Tentu saja penelitian ini menggunakan dataset dari analisis sentimen pandangan masyarakat Indonesia terhadap Capres nomor urut 01 yang belum pernah digunakan pada penelitian sebelumnya.

Parameswari & Prihandoko (2022) melakukan penelitian analisis sentimen dengan metode CNN, data tweet yang terkumpul kemudian diolah dan dianalisis sentimennya menggunakan metode CNN. Metode CNN diterapkan pada program sebagai metode klasifikasi. Pelatihan model menggunakan 8.975 data latih. Hasil uji coba model terhadap 3.051 data uji dengan epoch sebanyak 100 menunjukkan akurasi sebesar 86% dan pada penelitian yang dilakukan oleh Hidayatul Qudsi dkk. (2021) dengan metode serupa hasil yang didapatkan *DoubleMax* CNN dapat mengklasifikasi sentimen pada saran mahasiswa dengan sangat baik, yaitu mencapai Akurasi tertinggi sebesar 98%, *Recall* 98%, *Precision* 98% dan Skor F1 98%. Hasil ini menunjukkan bahwa CNN memiliki potensi yang bagus untuk diterapkan pada penelitian analisis sentimen Capres 01 kali ini.

CNN adalah salah satu bentuk jaringan saraf tiruan yang dapat mendeteksi informasi di berbagai posisi dengan akurasi yang sangat baik. Model ini telah memecahkan beberapa masalah dalam pemrosesan gambar dan pemrosesan bahasa alami otomatis seperti analisis opini, jawaban untuk pertanyaan, ringkasan teks, dan

lain-lain. Hal ini ditandai dengan arsitektur tertentu untuk memfasilitasi pembelajaran (Rhanoui dkk., 2019). CNN juga dapat digunakan untuk mengklasifikasikan data teks, memungkinkan aplikasi dalam analisis sentimen berbasis teks. CNN digunakan untuk analisis sentimen karena kemampuannya mengekstrak fitur secara otomatis dari input kompleks (Shiri, 2023). Selain itu, CNN mampu memuat keseluruhan skala informasi klasifikasi objek tanpa kehilangan keakuratannya (Dhika dkk., 2020). Namun, CNN juga memiliki beberapa kekurangan. Salah satunya adalah bahwa CNN memerlukan data yang besar dan kompleks untuk dilatih, sehingga dapat menjadi sulit untuk diterapkan pada data yang terbatas (Sartini, 2020). Selain itu, CNN juga memerlukan proses *preprocessing* yang kompleks untuk mengubah data menjadi format yang sesuai dengan CNN, seperti mengubah teks menjadi vektor melalui proses *word embedding*. Dalam beberapa kasus, CNN juga dapat menjadi *overfitting* jika tidak dilakukan dengan baik, yang dapat mengurangi akurasi analisis sentimen. Tujuan dari penelitian ini adalah untuk memperbaiki dan meningkatkan akurasi analisis sentimen dengan menggunakan metode CNN. Penelitian ini menghadirkan kebaruan dengan melakukan modifikasi CNN dengan penambahan lapisan *dropout* dan *batch Normalization* dan data yang sudah dibalancing.