

I. PENDAHULUAN

Indonesia adalah suatu negara yang menganut sistem demokrasi. Hal ini ditandai dengan diadakannya suatu pemilihan umum terhadap Kepala Pemerintahan ataupun Kepala Daerah ,salah satunya adalah Pemilihan Presiden dan Wakil Presiden (Nurhuda et al., 2016). Masyarakat Indonesia sangat menantikan momen politik tersebut. Seorang tokoh politik yang ingin mengajukan diri menjadi calon Presiden tentu akan mempertimbangkan popularitas mereka berdasarkan opini di masyarakat. Pada era modern seperti saat ini, masyarakat menggunakan social media untuk menyampaikan pendapat tentang suatu isu terutama tentang pemilu Presiden (Asmara et al., 2020). Masyarakat sering menggunakan media sosial Twitter, Pada tahun 2019, ada 500 juta tweet yang dikirim setiap hari di Twitter, sebagian besar digunakan untuk berbagi informasi dan mengungkapkan perasaan pengguna (*Kementerian Komunikasi Dan Informatika*, n.d.).

Dengan kemajuan teknologi saat ini, masyarakat bisa mendapatkan informasi dengan cepat, yang berdampak besar pada banyak aspek kehidupan (Akram & Kumar, 2017). Calon presiden dapat lebih dekat dengan masyarakat melalui media sosial seperti Twitter, yang memungkinkan mereka untuk berinteraksi langsung dengan orang-orang dan menyampaikan aspirasi mereka. Visi dan misi dapat dibagikan secara cepat melalui platform ini. Keprihatinan terhadap kebutuhan masyarakat ditunjukkan dengan respons cepat terhadap masalah terbaru. Interaksi di media sosial juga menyampaikan keluhan dan harapan masyarakat, yang mendorong perbaikan kebijakan. Calon presiden dapat menjangkau generasi muda yang akrab dengan teknologi digital dengan meningkatkan inklusi kampanye dan partisipasi politik melalui sesi tanya jawab langsung.

Oleh karena itu dapat dirancang sistem yang secara otomatis akan mengelompokkan opini dan review yang ada sesuai kelasnya. Kelas sentimen dibagi menjadi kelas sangat positif, positif, netral, negatif, dan sangat negatif sehingga pengguna dapat membaca memilih ingin membaca opini sesuai yang diinginkan (Gunawan et al., 2018). Analisis sentiment adalah alat yang sangat penting untuk memahami pendapat berupa data teks. Dalam konteks analisis sentiment ini, terdapat dua pendekatan yaitu *Machine Learning* dan *Deep Learning*. Teknologi *Machine Learning* merupakan alat yang dikembangkan untuk belajar sendiri tanpa arahan

penggunanya. Sedangkan *Deep Learning* merupakan sub bidang *Machine Learning* yang dapat menangani data yang kompleks lebih baik dari *Machine Learning*. Dalam analisis sentiment, *Deep learning* digunakan untuk melatih model yang dapat memahami data teks.

Penelitian mengenai Analisis Sentiment pada Twitter sudah banyak dilakukan, contohnya pada penelitian (Febriansyah et al., 2023) Analisis Sentiment di Twitter Terhadap Anies Baswedan sebagai Bakal Calon Presiden 2024 Menggunakan Metode K-Nearest Neighbor, mengevaluasi sentimen dengan metode K-Nearest Neighbor dengan akurasi 87,35%. Keunggulan dari penelitian ini adalah evaluasi akurasi dan kombinasi parameter nilai $k = 14$ dan $\text{threshold} = 12$ yang menghasilkan nilai akurasi tertinggi pada penelitian tersebut. Dalam hal ini, bukti menunjukkan bahwa algoritma K-Nearest Neighbor berhasil diterapkan dengan kombinasi nilai $k = 14$ dan nilai $\text{threshold} = 14$, dapat meningkatkan performa model K-Nearest Neighbor. Selain itu, penelitian lain menggunakan algoritma Long Short-Term Memory (LSTM), yang menghasilkan evaluasi akurasi sebesar 82% dalam menganalisis sentimen terhadap Prabowo Subianto. Algoritma ini unggul dibandingkan dengan algoritma K-Nearest Neighbor. Selain itu, (Wibowo, n.d.) menggunakan CNN untuk menganalisis sentiment dengan hasil akurasi 92,62% dan hasilnya dapat digunakan untuk menampilkan opini publik terkait dengan proses pemilihan umum. Selain itu, (Jurnal et al., 2022a) menggunakan BERT untuk menganalisis sentiment terhadap Undang-Undang Hukum Pidana Baru (KUHP) pada twitter. Peneliti menggunakan metode *fine-grained* sentiment analysis untuk membedakan sentiment yang lebih spesifik, seperti sentiment yang lebih negatif atau lebih positif, dengan hasil mencapai akurasi 81% pada data uji. Selain itu, (Farisi & Hadi, 2023) menggunakan *Recurrent Neural Network* untuk menganalisis sentiment di twitter terkait isu anies baswedan sebagai calon presiden 2024. Peneliti menggunakan RNN yang dikombinasikan dengan word embedding word2vec. Hasil pengujian mendapatkan akurasi yang cukup tinggi yaitu 86,27%.

RCNN pernah digunakan oleh He & Zhang (2019) untuk mengatasi masalah deteksi objek dalam lingkungan yang tidak terbatas (unrestricted environment) dengan menggunakan pendekatan domain adaptation dan transfer learning. Penelitian ini mengusulkan framework Multi-Adversarial Faster-RCNN (MAF) untuk deteksi objek yang tidak terbatas dengan meminimalkan perbedaan domain dalam representasi fitur. Framework MAF dievaluasi pada tugas deteksi objek

di beberapa dataset termasuk Cityscapes, KITTI, dan Sim10k. Hasil eksperimen menunjukkan bahwa MAF mencapai kinerja state-of-the-art, mengungguli detektor yang ada dalam tugas deteksi objek yang tidak terbatas. RCNN juga pernah digunakan oleh Q. Zhang et al. (2020) untuk mengembangkan algoritma segmentasi deteksi kerusakan kendaraan berbasis Improved Mask RCNN. Algoritma ini diharapkan dapat meningkatkan akurasi dan efisiensi deteksi serta segmentasi area kerusakan kendaraan akibat kecelakaan lalu lintas. Hasil eksperimen menunjukkan bahwa algoritma Improved Mask RCNN yang diusulkan memiliki nilai Average Precision (AP), akurasi deteksi, dan akurasi masking yang lebih baik dibandingkan dengan metode sebelumnya. Hal ini meningkatkan efisiensi dalam menyelesaikan masalah kompensasi kecelakaan lalu lintas. Dari dua penelitian diatas dapat dilihat bahwa model RCNN hasil yang cukup bagus dan memiliki potensi apabila dicoba untuk penelitian analisis sentimen Capres 02. Penelitian tentang analisis sentimen Capres 02 di media sosial tidak dapat diabaikan karena tidak ada pemahaman yang mendalam tentang persepsi dan opini masyarakat yang beredar secara online.

Imbalanced data merupakan salah satu kekurangan atau tantangan utama dalam analisis sentiment, karena data yang tidak seimbang dapat menyebabkan model belajar pada bagian data yang lebih banyak dan mengabaikan bagian data yang lebih sedikit. Hal ini dapat menyebabkan model yang tidak akurat dalam menganalisis sentiment. Untuk mengatasi masalah ini, beberapa teknik dapat digunakan, seperti oversampling, undersampling, dan SMOTE (Synthetic Minority Over-sampling Technique) (Tanimoto et al., 2022). Teknik-teknik ini dapat membantu dalam memperbaiki keseimbangan data dan meningkatkan akurasi model.

Penelitian ini bertujuan untuk meningkatkan nilai akurasi dari penelitian sebelumnya, kebaruan dibanding dengan penelitian sebelumnya adalah penggunaan algoritma RCNN dan beberapa modifikasi seperti penambahan dropout dan convolutional layers dengan data yang sudah dibalancing menggunakan metode Random Over-Sampling. Penelitian sebelumnya mengandalkan hanya satu algoritma atau metode yang mungkin tidak optimal dalam mengatasi masalah overfitting dan ketidakseimbangan data. Beberapa penelitian juga tidak mengeksplorasi secara mendalam penggunaan RCNN dalam kombinasi dengan teknik pembalancing data yang tepat. Penelitian ini diharapkan dapat mengatasi permasalahan tersebut dan memberikan hasil yang lebih akurat dan lebih baik pada dataset yang digunakan.