

Implementasi Metode CNN dalam Mengklasifikasi Berita Bohong Dan Ujaran Kebencian UU ITE Dataset

Regina Izza Aofkarina¹, Zaehol Fatah²

¹ Teknologi Informasi, Universitas Ibrahimy, Situbondo

² Sistem Informasi, Universitas Ibrahimy, Situbondo

Email : reginaizzaaufkarina23@gmail.com

Abstrak

Penelitian ini bertujuan untuk mengimplementasikan metode Convolutional Neural Network (CNN) dalam mengklasifikasikan berita bohong dan ujaran kebencian pada dataset UU ITE. Penelitian menggunakan pendekatan Knowledge Discovery in Databases (KDD) yang diimplementasikan melalui platform RapidMiner. Dataset yang digunakan berasal dari sumber terbuka Kaggle yang berisi konten digital berbahasa Indonesia terkait dengan pelanggaran UU ITE. Proses penelitian dimulai dengan tahap preprocessing data yang meliputi case folding, penghapusan karakter khusus, dan normalisasi teks. Implementasi CNN dilakukan dengan menggunakan operator Deep Learning di RapidMiner, dikonfigurasi dengan fungsi aktivasi Rectifier dan 10 epochs untuk proses pembelajaran. Evaluasi model dilakukan menggunakan metrik standar termasuk akurasi, precision, dan recall. Hasil penelitian menunjukkan bahwa model CNN mencapai akurasi sebesar 49.99%, weighted mean recall 33.33%, dan weighted mean precision 16.67%. Confusion matrix menunjukkan 3684 data terklasifikasi sebagai True Positive dan 1842 data sebagai *False negative*. Meskipun hasil ini mengindikasikan perlunya pengembangan lebih lanjut, penelitian ini berhasil membangun fondasi untuk sistem klasifikasi otomatis konten digital berbahasa Indonesia. Rekomendasi untuk pengembangan mencakup peningkatan kualitas dataset, optimasi parameter model, dan penerapan teknik preprocessing tambahan untuk meningkatkan akurasi klasifikasi.

Kata Kunci: *Jaringan Syaraf Tiruan, Twitter, Ujaran Kebencian, Berita Palsu, RapidMiner*

PENDAHULUAN

Perkembangan teknologi informasi yang semakin pesat telah mengubah cara masyarakat dalam mengakses dan menyebarkan informasi. Media sosial dan platform digital menjadi sarana utama pertukaran informasi yang memungkinkan siapa saja dapat dengan mudah membagikan berita atau konten (Nursingah et al., 2024). Namun, kemudahan ini juga membawa tantangan baru berupa meningkatnya penyebaran berita bohong (hoax) dan ujaran kebencian yang dapat memecah belah masyarakat (Imaroh et al., 2023).

Di Indonesia, fenomena berita bohong dan ujaran kebencian menjadi masalah serius yang perlu mendapat perhatian khusus. Berdasarkan data Kementerian Komunikasi dan Informatika, sepanjang tahun 2020 tercatat lebih dari

1.200 konten hoax yang tersebar di berbagai platform digital (Romadlon, 2020). Kondisi ini diperparah dengan rendahnya tingkat literasi digital masyarakat Indonesia, di mana berdasarkan survei nasional literasi digital 2020 menunjukkan bahwa indeks literasi digital Indonesia masih berada pada angka 3.47 dari skala 5 (Suherdi et al., 2021).

Salah satu upaya untuk mengatasi permasalahan ini adalah dengan memanfaatkan teknologi kecerdasan buatan, khususnya dalam bidang pembelajaran mesin (machine learning). Convolutional Neural Network (CNN) sebagai salah satu metode deep learning telah menunjukkan keberhasilan dalam berbagai tugas pemrosesan bahasa alami, termasuk klasifikasi teks (Naquitasia et al., 2022). Kemampuan CNN dalam menangkap pola-pola kompleks dalam data

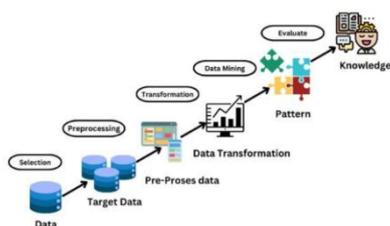
teks menjadikannya potensial untuk diterapkan dalam klasifikasi berita bohong dan ujaran kebencian.

Undang-Undang Informasi dan Transaksi Elektronik (UU ITE) yang berlaku di Indonesia telah mengatur tentang penyebaran berita bohong dan ujaran kebencian. Namun, proses identifikasi dan klasifikasi konten-konten tersebut masih banyak dilakukan secara manual, yang membutuhkan waktu dan sumber daya yang tidak sedikit. Implementasi sistem otomatis berbasis kecerdasan buatan dapat menjadi solusi untuk mempercepat proses identifikasi dan klasifikasi konten-konten bermasalah tersebut(MAHFI, 2021).

Penelitian ini bertujuan untuk mengembangkan model klasifikasi otomatis menggunakan metode CNN untuk mengidentifikasi berita bohong dan ujaran kebencian berdasarkan dataset UU ITE. Dengan menggunakan metodologi Knowledge Discovery in Databases (KDD), penelitian ini diharapkan dapat menghasilkan model yang mampu membantu proses identifikasi konten bermasalah secara lebih efektif dan efisien. Hasil dari penelitian ini dapat menjadi langkah awal dalam pengembangan sistem deteksi otomatis yang dapat membantu upaya penegakan UU ITE dan meningkatkan kualitas informasi yang beredar di masyarakat Indonesia(Arhami, 2024).

METODE

Metodologi penelitian dalam implementasi CNN untuk klasifikasi berita bohong dan ujaran kebencian ini terdiri dari empat tahapan utama. Setiap tahapan dirancang secara sistematis untuk memastikan hasil yang optimal dan dapat divalidasi(Alrasyid et al., 2024).



Gambar 1. Ilustrasi Proses Data Mining

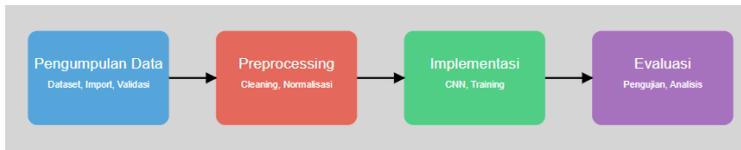
Tahap pertama adalah pengumpulan data, di mana dataset UU ITE yang berisi berita bohong dan ujaran kebencian dikumpulkan. Pada tahap ini, data dikumpulkan dari berbagai sumber yang relevan dan diimpor ke dalam format yang sesuai untuk pengolahan. Dataset kemudian divalidasi untuk memastikan kelengkapan dan kesesuaian dengan kebutuhan penelitian. Proses ini meliputi pemeriksaan format data, pengecekan missing values, dan verifikasi label pada data.

Tahap kedua adalah preprocessing data, yang merupakan tahapan krusial dalam pengolahan teks. Pada tahap ini, dilakukan serangkaian proses untuk membersihkan dan mempersiapkan data, meliputi: case folding untuk menyeragamkan huruf menjadi lowercase, removal of special characters untuk menghilangkan karakter-karakter khusus, tokenization untuk memecah teks menjadi kata-kata individual, dan stopword removal untuk menghilangkan kata-kata yang tidak memiliki makna signifikan. Selain itu, dilakukan juga normalisasi teks untuk menstandarisasi kata-kata tidak baku menjadi baku(Jalil et al., 2024).

Tahap ketiga adalah implementasi, di mana model CNN dirancang dan dilatih menggunakan data yang telah dipreprocessing. Pada tahap ini, arsitektur CNN dikonfigurasi dengan mempertimbangkan berbagai parameter seperti jumlah layer, fungsi aktivasi, dan hyperparameter lainnya. Proses training dilakukan dengan membagi dataset menjadi data training dan testing dengan rasio yang sesuai. Model CNN kemudian dilatih menggunakan data training dengan epoch dan batch size yang telah ditentukan.

Tahap terakhir adalah evaluasi, di mana performa model CNN diuji dan dianalisis. Evaluasi dilakukan menggunakan berbagai metrik seperti accuracy, precision dan recall. Hasil evaluasi kemudian dianalisis untuk mengetahui efektivitas model dalam mengklasifikasi berita bohong dan ujaran kebencian(Nasta'in et al., 2024). Pada tahap

ini juga dilakukan analisis error untuk mengidentifikasi pola-pola kesalahan klasifikasi dan potensi perbaikan model.



Gambar 2. Metodologi Penelitian

HASIL DAN PEMBAHASAN

Preprocessing data

Dalam penelitian ini, data mentah awalnya diambil dari platform Kaggle dalam format CSV yang berisi kumpulan teks berita dan label klasifikasinya. Data tersebut kemudian dikonversi ke format Excel untuk memudahkan proses pengolahan dan visualisasi data awal. Dataset yang digunakan merupakan kumpulan teks berbahasa Indonesia yang terkait dengan berita bohong dan ujaran kebencian yang termasuk dalam ranah UU ITE.



Gambar 3. Dataset yang digunakan pada Kaggle

Tahap preprocessing dimulai dengan pembersihan data menggunakan operator Replace di RapidMiner. Proses ini melibatkan beberapa langkah penting. Pertama, dilakukan case folding untuk mengubah semua teks menjadi huruf kecil (lowercase) agar standar. Misalnya, kata "Berita", "BERITA", dan "berita" akan diubah menjadi format yang sama yaitu "berita". Selanjutnya, dilakukan penghapusan karakter-karakter yang tidak diperlukan seperti tanda baca, angka, dan simbol-simbol khusus menggunakan regular expression (regex). URL, mention (@username), hashtag (#), dan emoticon juga dihapus karena tidak memberikan kontribusi signifikan terhadap proses klasifikasi. Proses ini penting untuk

mengurangi noise dalam data dan memfokuskan analisis pada konten teks yang substantif.

Setelah pembersihan dasar, dilakukan tokenisasi untuk memecah teks menjadi kata-kata individual. Pada tahap ini juga diterapkan stopwords removal untuk menghilangkan kata-kata umum dalam bahasa Indonesia yang tidak memiliki makna khusus seperti "yang", "di", "ke", "dari", dan sebagainya. Proses ini menggunakan daftar stopwords bahasa Indonesia yang telah distandarisasi. Tahap terakhir preprocessing adalah normalisasi teks, di mana kata-kata tidak baku atau singkatan diubah menjadi bentuk bakunya. Misalnya "gk" menjadi "tidak", "yg" menjadi "yang", dan "dgn" menjadi "dengan". Proses ini penting mengingat banyaknya penggunaan bahasa informal dan singkatan dalam konten media sosial dan berita online.



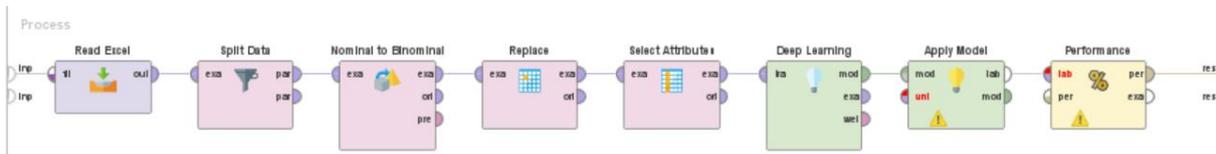
Gambar 4. Dataset telah convert Excel

Implementasi K-Means Clustering

Dalam implementasinya, model CNN dikonfigurasi dengan beberapa parameter utama. Parameter activation function menggunakan Rectifier (ReLU), yang dipilih karena kemampuannya dalam menangani non-linearitas dan mencegah masalah vanishing gradient pada proses pembelajaran. Jumlah epochs diatur pada nilai 10.0, yang menentukan berapa kali model akan melakukan iterasi pembelajaran pada keseluruhan dataset training.

Pada operator Deep Learning, hidden layer sizes dikonfigurasi dengan menggunakan Edit Enumeration untuk menyesuaikan arsitektur jaringan dengan kompleksitas data. Parameter ini

menentukan jumlah node pada setiap hidden layer dalam jaringan. Pengaturan ini penting untuk memastikan model memiliki kapasitas yang cukup untuk mempelajari pola-pola dalam data teks yang kompleks.



Gambar 5. Tahapan Proses RapidMiner

Proses training model menggunakan data yang telah melalui tahap preprocessing, di mana data teks telah diubah menjadi format yang sesuai untuk input CNN. Output dari operator Deep Learning kemudian dihubungkan ke operator Apply Model untuk melakukan prediksi pada data testing. Model yang telah dilatih akan menghasilkan klasifikasi untuk setiap teks input ke dalam kategori yang telah ditentukan.

Evaluasi Model

Evaluasi model CNN dalam penelitian ini dilakukan dengan menggunakan operator Performance di RapidMiner yang menghasilkan beberapa metrik penting untuk mengukur kinerja model. Berdasarkan hasil pengujian yang telah dilakukan, model mencapai tingkat akurasi sebesar 49.99%, yang mengindikasikan bahwa model mampu mengklasifikasikan setengah dari total data dengan benar.

Dari hasil confusion matrix yang dihasilkan, dapat dilihat bahwa model memiliki distribusi prediksi yang menunjukkan 3684 data terklasifikasi sebagai True Positive (benar positif) dan 1842 data sebagai False negative (salah negatif). Nilai weighted mean recall yang mencapai 33.33% menunjukkan kemampuan model dalam mengidentifikasi kasus positif yang sebenarnya, sementara weighted mean precision sebesar 16.67% mengindikasikan tingkat ketepatan model dalam melakukan klasifikasi positif.

Performa model yang belum optimal ini kemungkinan dipengaruhi oleh

beberapa faktor. Pertama, kompleksitas bahasa Indonesia dalam konteks berita bohong dan ujaran kebencian yang memiliki banyak variasi dan nuansa. Kedua, kemungkinan adanya ketidakseimbangan dalam dataset yang digunakan untuk training. Ketiga, parameter CNN yang mungkin masih memerlukan penyesuaian lebih lanjut untuk meningkatkan performa model.

PerformanceVector

```
PerformanceVector:
accuracy: 49.99%
ConfusionMatrix:
True:  Yes  No  0
Yes:  2947 1474 1474
No:   0    0    0
0:    0    0    0
weighted_mean_recall: 33.33%, weights: 1, 1, 1
ConfusionMatrix:
True:  Yes  No  0
Yes:  2947 1474 1474
No:   0    0    0
0:    0    0    0
weighted_mean_precision: 16.66%, weights: 1, 1, 1
ConfusionMatrix:
True:  Yes  No  0
Yes:  2947 1474 1474
No:   0    0    0
0:    0    0    0
```

Gambar 6. Tahapan Proses RapidMiner

KESIMPULAN

Berdasarkan penelitian yang telah dilakukan tentang implementasi metode CNN dalam mengklasifikasikan berita bohong dan ujaran kebencian pada dataset UU ITE menggunakan RapidMiner dengan pendekatan KDD, diperoleh beberapa hasil penting. Model CNN yang dikembangkan menunjukkan performa dengan akurasi sebesar 49.99%, weighted mean recall 33.33%, dan weighted mean precision 16.67%. Hasil ini menunjukkan bahwa model telah mampu melakukan klasifikasi

dasar meskipun masih memerlukan pengembangan lebih lanjut untuk meningkatkan akurasi.

Proses preprocessing data yang meliputi case folding, penghapusan karakter khusus, dan normalisasi teks telah berhasil mempersiapkan data untuk proses klasifikasi. Implementasi CNN dengan konfigurasi fungsi aktivasi Rectifier dan 10 epochs menunjukkan kemampuan model dalam mempelajari pola-pola dalam teks berbahasa Indonesia. Hasil confusion matrix menunjukkan bahwa model memiliki kemampuan yang cukup dalam mengidentifikasi kasus positif dengan 3684 data terklasifikasi sebagai True Positive, meskipun masih terdapat 1842 data yang terklasifikasi sebagai *False negative*.

Berdasarkan hasil penelitian ini, terdapat beberapa saran untuk pengembangan lebih lanjut. Pertama, diperlukan peningkatan kualitas dataset dengan cara memperbesar jumlah data training dan melakukan balancing dataset untuk mengatasi ketidakseimbangan kelas yang dapat mempengaruhi performa model. Kedua, optimasi model dapat dilakukan dengan melakukan eksperimen menggunakan arsitektur CNN yang berbeda dan mencoba variasi parameter seperti learning rate, jumlah epoch, dan ukuran batch untuk menemukan konfigurasi optimal. Ketiga, penerapan teknik preprocessing tambahan seperti lemmatization dan stemming dapat dipertimbangkan untuk meningkatkan kualitas input data, serta mengembangkan pendekatan hybrid dengan menggabungkan CNN dengan metode lain seperti LSTM untuk meningkatkan akurasi klasifikasi.

DAFTAR PUSTAKA

Alrasyid, H., Homaidi, A., Kom, M., Fatah, Z., & Kom, M. (2024). *Comparison Support Vector Machine and Random Forest Algorithms in Detect Diabetes*. *I(1)*, 447–453.

Arhami, M. (2024). *Metodologi Penelitian Untuk Teknologi Informasi Dan Komputer*.

Imaroh, Z., Irawan, A. A., & Dkk. (2023). Pertanggungjawaban Pidana Penyebaran Berita Hoax Di Media Sosial. In *Hukum*.

Jalil, A., Homaidi, A., & Fatah, Z. (2024). Implementasi Algoritma Support Vector Machine Untuk Klasifikasi Status Stunting Pada Balita. *G-Tech: Jurnal Teknologi Terapan*, *8(3)*, 2070–2079. <https://doi.org/10.33379/gtech.v8i3.4811>

MAHFI, M. R. R. (2021). Undang-Undang Informasi Dan Transaksi Elektronik (Uu Ite) Dalam Perspektif Hukum Pidana Administrasi (Administrative Penal Law). *Badamai Law Journal*, *5(1)*, 140. <https://doi.org/10.32801/damai.v5i1.10055>

Naquitasia, R., Fudholi, D. H., & Iswari, L. (2022). *ANALISIS SENTIMEN BERBASIS ASPEK PADA WISATA HALAL DENGAN METODE DEEP LEARNING*. *16*, 156–164.

Nasta'in, M., Munazilin, A., & Susanto, A. (2024). Predict The Trend Of Indonesian People's Interest In Bitcoin Facing The 2024 Bitcoin Halving Using A Forecasting Algorithm. *Journal Ibrahi*, *1(1)*, 463–468.

Nursingah, L., Ruuhwan, R., & Mufizar, T. (2024). Analisis Sentimen Pengguna Aplikasi X Terhadap Program Makan Siang Gratis Dengan Metode Naïve Bayes Classifier. *Jurnal Informatika Dan Teknik Elektro Terapan*, *12(3)*. <https://doi.org/10.23960/jitet.v12i3.4336>

Romadlon, F. (2020). COVID-19 Dalam Ragam Tinjauan Perspektif. In *COVID-19 Dalam Ragam Tinjauan Perspektif*.

Suherdi, D., Rezky, S. F., Apdilah, D., Sinuraya, J., Sahputra, A., Syahputra, D., & Wahyuni, D. (2021). Peran Literasi Digital Di Masa Pandemi. In *Cattleya Darmaya Fortuna* (Vol. 123). <https://is.gd/q7Wy1w>