

Machine Learning Untuk Deteksi Berita Hoax Menggunakan BERT

Isnaeni Imroatus Sholikhah*¹, Aris Tri Jaka Harjanta², Khoiriya Latifah³

^{1,2,3}Program Studi Informatika, Universitas PGRI Semarang, Kota Semarang

*Email korespondensi: isnaeniimroatussholikhah30@gmail.com

Abstract.

The spread of hoax news is a serious problem in a digital era filled with information that is easily spread. This research aims to overcome these challenges by using BERT (Bidirectional Encoder Representations from Transformers) and Random Forest Classifier approaches in modeling and classifying hoax news. The BERT method is used to generate a rich sentence embedding representation, while the Random Forest Classifier is used as an effective classification tool and is able to overcome the problem of overfitting on text data. Using Google Collaboratory as a tool to explore the dataset, the evaluation results show that the BERT classification model has an accuracy of 67 percent on validation data in classifying hoax news, which shows the performance of the Machine Learning model in classifying hoax news. F1-Score for positive labels is 0.67 with Precision 0.57 for positive labels, Recall for positive labels is 0.80, the potential of this method can help in combating the spread of fake news. This research makes an important contribution in an effort to build a system that can identify and combat the spread of fake news, as well as show how to use BERT in the analysis of hoax news classification.

Keywords: Machine Learning, hoax news classification, BERT, Random Forest Classifier, embedding representation.

Abstrak

Penyebaran berita hoax menjadi permasalahan serius dalam era digital yang dipenuhi dengan informasi yang mudah tersebar. Penelitian ini bertujuan untuk mengatasi tantangan tersebut dengan menggunakan pendekatan *BERT (Bidirectional Encoder Representations from Transformers)* dan *Random Forest Classifier* dalam memodelkan dan mengklasifikasikan berita hoax. Metode *BERT* digunakan untuk menghasilkan representasi embedding kalimat yang kaya, sementara *Random Forest Classifier* digunakan sebagai alat klasifikasi yang efektif dan mampu mengatasi masalah overfitting pada data teks. Dengan Menggunakan *Google Colaboratory* sebagai alat untuk mengeksplor dataset sehingga hasil evaluasi menunjukkan bahwa model klasifikasi *BERT* memiliki akurasi sebesar 67 Persen pada data validasi dalam mengklasifikasikan berita hoax, yang menunjukkan performance atau kinerja model *Machine Learning* dalam melakukan klasifikasi berita hoax. *F1-Score* untuk label positif adalah 0.67 dengan Presisi 0.57 untuk label positif, Recall untuk label positif adalah 0.80, potensi metode ini dapat membantu dalam memerangi penyebaran berita palsu. Penelitian ini memberikan kontribusi penting dalam upaya membangun sistem yang dapat mengidentifikasi dan memerangi penyebaran berita palsu, serta menunjukkan cara penggunaan *BERT* dalam analisis klasifikasi berita hoax.

Kata Kunci: Machine Learning, Klasifikasi berita hoax, BERT, Random Forest Classifier, representasi embedding.

1. Pendahuluan

Dalam bidang penelitian ilmiah, makalah adalah salah satu sarana penting bagi peneliti untuk menganalisis dan memahami bidang ini. Namun, seiring dengan perkembangan zaman, jenis jurnal dan jumlah makalah semakin meningkat. Sangat penting untuk memanfaatkan teknologi yang sedang berkembang untuk menambang dan menganalisis topik-topik hangat di bidang ilmiah, untuk membantu peneliti dalam menganalisis topik-topik hangat di sekitar, untuk memberikan dasar bagi para peneliti di bidang tertentu. Pemerintah dan perusahaan dapat juga dapat fokus pada penelitian di berbagai sektor, dari mana mereka dapat menemukan yang terbaru arah penelitian terbaru dan menangkap peluang pertama [1]. Penelitian karya ilmiah yang ada saat ini terutama menggunakan teknologi seperti analisis kata kunci, analisis kutipan, dan analisis topik untuk menambang topik berita, sambil menggunakan evolusi topik untuk menganalisis perubahan dan untuk mendemonstrasikan proses fenomena seperti penipuan dalam informasi dan berita. Namun, penelitian yang ada tidak menghilangkan gangguan teks tepi selama data pra-pemrosesan data, serta penambangan topik dengan sedikit analisis semantik teks konteks teks, memperlakukan teks sebagai kumpulan kata-kata saja. Saat menyaring asosiasi topik, seringkali didasarkan pada pengalaman peneliti, dan aturan pemfilteran asosiasi subjek adalah jarang digunakan. Bertujuan untuk mengatasi kekurangan dari penelitian ini, makalah ini pertama-tama menggunakan kalimat transformer untuk menghasilkan set teks [2]. Kemudian BERT [3] digunakan untuk mempelajari kontekstual kontekstual dari teks. Akhirnya, aturan penyaringan asosiasi topik dibuat untuk menghapus asosiasi topik yang tidak penting dan membangun jalur evolusi.

Salah satu masalah yang perlu dipecahkan dalam bidang ini adalah klasifikasi berita hoaks dengan akurasi yang tinggi. Dalam era penyebaran informasi palsu yang semakin meningkat, penting untuk mengembangkan metode yang efektif dalam mengidentifikasi dan membedakan antara artikel berita asli dan berita hoaks. Sebelumnya, penelitian belum sepenuhnya mengeksplorasi potensi teknik pemodelan topik dengan menggunakan model bahasa kontekstual seperti BERT untuk meningkatkan akurasi klasifikasi berita hoaks. Integrasi model ini dapat meningkatkan kekuatan metode tersebut dan memberikan pemahaman yang lebih komprehensif tentang konten berita.

Selain BERT, salah satu metode yang dapat digunakan dalam klasifikasi berita hoaks adalah Random Forest.[4] Random Forest adalah algoritma pembelajaran mesin yang kuat dan dapat digunakan untuk membangun model klasifikasi berbasis pohon keputusan. Metode ini memanfaatkan keputusan yang diambil dari sejumlah pohon keputusan acak yang bekerja bersama-sama untuk menghasilkan prediksi akhir. Dengan menggunakan Random Forest, dapat dilakukan analisis yang lebih mendalam pada fitur-fitur dan hubungan antara data untuk menghasilkan prediksi yang akurat. Dalam konteks klasifikasi berita hoaks, Random Forest dapat mengambil fitur-fitur seperti kata-kata kunci, struktur kalimat, entitas yang disebutkan, dan fitur linguistik lainnya untuk membedakan antara berita hoax dan non-hoax.

2. Metode

1. Machine Learning

Machine Learning adalah pengembangan sistem yang dapat belajar dan beradaptasi dari data. Pembelajaran mesin merupakan bagian dari kecerdasan buatan yang banyak digunakan untuk memecahkan berbagai masalah. [5] Salah satu pendekatan yang dapat digunakan dalam pengembangan dan penerapan Machine Learning adalah CRISP-DM (Cross-Industry Standard Process for Data Mining).

CRISP-DM adalah metodologi yang terstruktur dan teruji waktu yang digunakan dalam proses pengembangan proyek data mining, termasuk dalam konteks Machine Learning. Meskipun CRISP-DM dirancang khusus untuk data mining, namun prinsip dan tahapannya di sajikan di gambar 1. Dapat dengan mudah diterapkan dalam konteks Machine Learning.



Gambar 1. Tahapan CRISP-DM

Sumber <https://binus.ac.id/malang/2022/05/crisp-dm-cross-industry-standard-process-for-data-mining/>

Tujuannya adalah untuk memungkinkan komputer untuk melakukan tugas-tugas tertentu tanpa perlu diprogram secara eksplisit. Dalam Machine Learning, komputer diajarkan untuk mengenali pola dan membuat prediksi atau keputusan berdasarkan data yang diberikan. Ada beberapa pendekatan dalam Machine Learning, termasuk supervised learning, unsupervised learning, dan reinforcement learning.

2. Klasifikasi Teks Topic Modeling BERT

BERT (Bidirectional Transformers) [2] secara lengkap adalah Representasi Encoder Dua Arah dari Transformers . Ini adalah model penyandian dua arah yang khas yang dibangun dengan transformator arsitektur transformer [6], yang terdiri dari beberapa encoder dan decoder.[7] Model tersebut mengemulasi tugas tes cloze untuk menyelesaikan tugas tingkat token dengan mengambil MLM (Masked Language Model)) untuk secara acak menutupi sebuah kata dalam teks dan menggunakan mekanisme perhatian dua arah yang lebih baik untuk memprediksi kata menggunakan kalimat sebelum dan sesudahnya, sehingga membuat model yang telah dilatih untuk representasi dua arah. Untuk mengimplementasikan representasi tingkat kalimat, prediksi kalimat berikutnya dilakukan dengan menggunakan metode NSP (Next Sentence Prediction).[8] Menerapkan BERT untuk klasifikasi teks dan mengusulkannya sebagai ekstraktor fitur, di mana vektor yang dihasilkan vektor yang dihasilkan hanya perlu dihubungkan dengan lapisan output untuk mencapai hasil yang diinginkan dalam berbagai tugas hilir. Selain itu, model yang dikembangkan berdasarkan BERT meliputi RoBERTa, yang dilatih pada korpus yang lebih besar ALBERT [9], yang mengurangi jumlah parameter dan BERT , yang memperkenalkan topik informasi.

3. Hasil dan Pembahasan

3.1. Penyajian Hasil

Hasil dan diskusi dapat dilakukan secara keseluruhan yang berisi temuan dan penjelasan penelitian.

1. Topic BERT

Model Klasifikasi Berikut ini adalah langkah dalam pemrosesan teks untuk Klasifikasi Berita Hoax Menggunakan Model BERT:

1. Preprocessing pada dataset: Ini melibatkan pembersihan teks, penghapusan tanda baca atau karakter khusus, konversi ke huruf kecil, dan langkah-langkah lainnya untuk mempersiapkan data teks.
2. Mempersiapkan input untuk model BERT: Ini mencakup tokenisasi teks menggunakan tokenizer BERT, membatasi panjang teks, dan mengubah token menjadi input IDs dan attention masks yang dibutuhkan oleh model BERT.
3. Menghitung embedding menggunakan BERT: Menggunakan model BERT yang telah diinisialisasi sebelumnya, teks yang telah di-tokenisasi dikirimkan ke model untuk menghitung embeddingnya. Dalam kasus ini, embedding diambil dari state tersembunyi terakhir token [CLS].
4. Membagi data menjadi data pelatihan dan data validasi: Data yang telah di-preprocess dan embedding yang dihitung dibagi menjadi set pelatihan dan set validasi untuk melatih dan menguji model klasifikasi.
5. Melatih model klasifikasi: Menggunakan set pelatihan, model klasifikasi (dalam kasus ini, Random Forest Classifier) dilatih pada embedding BERT dan label yang sesuai.
6. Melakukan prediksi pada data validasi: Setelah melatih model, dilakukan prediksi pada set validasi menggunakan embedding yang dihitung sebelumnya.
7. Menampilkan laporan klasifikasi dan matriks kebingungan: Hasil prediksi dibandingkan dengan label asli dari set validasi, dan kemudian dilaporkan metrik klasifikasi seperti presisi, recall, skor F1, dan matriks kebingungan untuk mengevaluasi kinerja model.
8. Melakukan prediksi pada data uji.

Metode ini menggunakan model BERT untuk melakukan klasifikasi berita hoax. Model klasifikasi yang digunakan pada contoh ini adalah Random Forest, namun dapat diganti dengan model klasifikasi lain sesuai kebutuhan.

2. Random Forest

Random Forest adalah algoritma pembelajaran mesin yang kuat dan dapat digunakan untuk membangun model klasifikasi berbasis pohon keputusan.

- Laporan Klasifikasi (Classification Report):

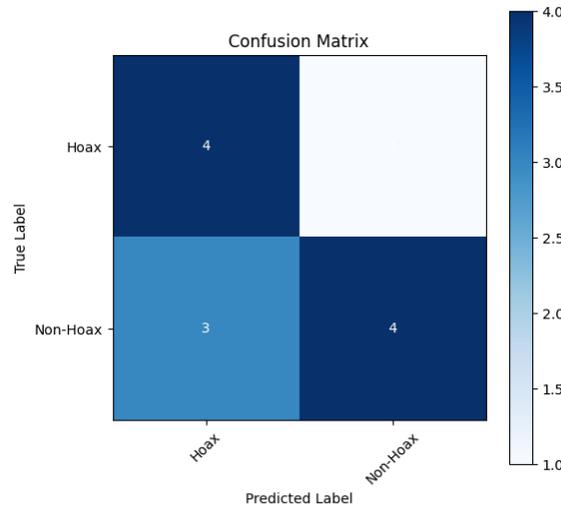
Laporan klasifikasi menyediakan matrik evaluasi yang lebih rinci untuk masing-masing kelas dalam masalah klasifikasi. Metrik yang umum digunakan dalam laporan klasifikasi meliputi presisi (precision), recall, f1-score, dan akurasi (accuracy).

- Matriks Kebingungan (Confusion Matrix):

Matriks kebingungan adalah tabel yang menggambarkan klasifikasi yang benar dan klasifikasi yang salah dari model. Matriks ini membantu dalam mengevaluasi kinerja model dengan melihat jumlah data yang diklasifikasikan dengan benar dan salah untuk setiap kelas.

Dengan menganalisis laporan klasifikasi dan matriks kebingungan, kita dapat mendapatkan pemahaman yang lebih baik tentang bagaimana model Random Forest

bekerja dalam mengklasifikasikan data. Dalam konteks klasifikasi berita hoax, kita dapat melihat seberapa baik model dapat membedakan antara berita hoax dan bukan hoax. Berikut hasil Klasifikasi disajikan di gambar. 1.



Gambar 2. Gambar hasil klasifikasi

3.2. Pembahasan

Dari Hasil yang sudah didapatkan dari penelitian, berikut tabel 1. dari data Dummy yang diambil beberapa sample saja:

Tabel 1. Tabel Sample Dataset Dummy dan Label

Dataset	Label
Ini adalah berita hoax	1
Berita ini tidak benar	0
Berita hoax sering menyesatkan	1
Berita ini adalah kebohongan	0
Hoax adalah berita palsu	1

1. F1-score

adalah ukuran gabungan dari presisi dan recall. F1-score menggabungkan presisi dan recall untuk memberikan nilai tunggal yang mencerminkan performa keseluruhan model.

Rumus: $F1-Score = 2 * (Precision * Recall) / (Precision + Recall)$

Dalam rumus di atas:

- TP adalah jumlah data positif yang diklasifikasikan dengan benar sebagai positif (hoax).
- FP adalah jumlah data negatif yang salah diklasifikasikan sebagai positif (hoax).
- FN adalah jumlah data positif yang salah diklasifikasikan sebagai negatif (non-hoax).

2. Precision.

Precision mengukur sejauh mana model dapat mengenali dengan benar kelas positif (hoax) dari semua hasil yang diklasifikasikan sebagai positif.[10] ada di bintang/link

Rumus: $Precision = TP / (TP + FP)$

True Positive (TP) : Jumlah data yang bernilai Positif dan diprediksi benar sebagai Positif.

False Positive (FP) : Jumlah data yang bernilai Negatif tetapi diprediksi sebagai Positif.

3. Recall.

Recall, juga dikenal sebagai sensitivitas atau true positive rate, mengukur sejauh mana model dapat mengenali dengan benar semua sampel positif (hoax).

Rumus: $Recall = TP / (TP + FN)$

FN = False Negative (jumlah data positif yang salah diklasifikasikan sebagai negatif)

Yang menunjukkan tabel hasil penelitian yang baik disajikan pada Tabel 2 sebagai berikut.

Tabel 2. Tabel hasil penelitian kinerja model

	Precision	Recall	F1-Score	Support
0	0.57	0.80	0.67	5
1	0.80	0.57	0.67	7
Accuracy			0.67	12

Matrik evaluasi ini digunakan untuk mengevaluasi performa model klasifikasi dalam mengenali berita hoax. Semakin tinggi presisi, recall, dan f1-score, semakin baik performa model dalam mengklasifikasikan berita hoax dengan benar.[10]

Pada bagian pembahasan, hasil dari penelitian tersebut diinterpretasikan dan dianalisis. Berikut adalah pembahasan yang menjelaskan hasil penelitian:

Pada hasil penelitian terbaru, dilakukan klasifikasi menggunakan model BERT untuk memprediksi label positif (1) atau negatif (0) dari suatu teks. Setelah melatih model klasifikasi dengan menggunakan data pelatihan, dilakukan evaluasi pada data validasi untuk menganalisis kinerja model. Hasil evaluasi menunjukkan bahwa model klasifikasi BERT memiliki akurasi sebesar 67% pada data validasi. Selain akurasi, juga diperhatikan precision, recall, dan f1-score. Precision mengukur sejauh mana prediksi positif yang benar, recall mengukur sejauh mana kemampuan model dalam menemukan kasus positif, dan f1-score adalah ukuran harmonis antara precision dan recall.

Dalam kasus ini, model Random Forest memiliki precision sebesar 0.57 untuk label positif. Artinya, sekitar 57% dari prediksi positif yang dibuat oleh model adalah benar-benar positif. Recall untuk label positif adalah 0.80, yang berarti model berhasil menemukan sekitar 80% dari keseluruhan kasus positif yang ada. F1-Score untuk label positif adalah 0.67, yang merupakan ukuran rata-rata dari precision dan recall

Matriks kebingungan (confusion matrix) juga memberikan gambaran tentang hasil klasifikasi. Dalam matriks kebingungan ini, terdapat empat kategori: true positive (TP), false positive (FP), true negative (TN), dan false negative (FN). Dalam kasus ini, terdapat 4 TP, yang berarti ada 4 prediksi benar untuk kasus positif. Terdapat juga 1 FP, yang merupakan prediksi salah untuk kasus negatif yang sebenarnya positif. Selain itu, terdapat 3 TN, yang berarti ada 3 prediksi benar untuk kasus negatif, dan 3 FN, yang merupakan prediksi salah untuk kasus positif yang sebenarnya negatif. Dengan melihat hasil ini, kita dapat menyimpulkan bahwa model Random Forest memiliki performa yang cukup baik dalam mengklasifikasikan berita hoax dengan akurasi sebesar 0.67.

Dari Hasil Prediksi data uji dari 5 sample data dummy diatas menghasilkan

Teks: Ini adalah berita hoax
Prediksi: Hoax

Teks: Berita ini tidak benar
Prediksi: Non-Hoax

Teks: Berita hoax sering menyesatkan
Prediksi: Hoax

Teks: Berita ini adalah kebohongan
Prediksi: Non-Hoax

Teks: Hoax adalah berita palsu
Prediksi: Hoax

4. Kesimpulan

Berdasarkan hasil penelitian yang dilakukan menggunakan model BERT untuk klasifikasi teks, berikut adalah kesimpulan yang dapat diambil:

1. Model BERT memiliki kemampuan yang baik dalam memahami dan menganalisis teks, sehingga cocok digunakan untuk tugas klasifikasi berita hoax.
2. Dalam penelitian ini, penggunaan representasi BERT sebagai fitur input untuk model klasifikasi Random Forest menghasilkan akurasi sebesar 0.67.
3. Precision untuk label positif adalah 0.57, yang berarti sekitar 57% dari prediksi positif model benar-benar positif.
4. Recall untuk label positif adalah 0.80, yang berarti model berhasil menemukan sekitar 80% dari keseluruhan kasus positif yang ada.
5. F1-Score untuk label positif adalah 0.67, yang merupakan ukuran rata-rata dari precision dan recall.
6. Meskipun model telah memberikan hasil yang cukup baik, masih terdapat ruang untuk meningkatkan performa dengan mengoptimalkan parameter atau menggunakan metode lain yang lebih cocok.
7. Hasil matriks kebingungan (confusion matrix) menunjukkan bahwa terdapat prediksi yang benar dan salah baik untuk kasus positif maupun negatif.
8. Kesimpulan ini didasarkan pada dataset dan kondisi yang digunakan dalam penelitian ini, sehingga hasilnya mungkin berbeda jika digunakan pada dataset atau kondisi yang berbeda.

Untuk melakukan peningkatan penelitian selanjutnya, disarankan untuk melakukan eksplorasi lebih lanjut tentang bagaimana meningkatkan kinerja model BERT dalam kasus klasifikasi teks yang lebih rumit. Selain itu, dapat juga dipertimbangkan untuk menggunakan teknik pemrosesan teks tambahan atau memperluas dataset untuk meningkatkan generalisasi model.

5. Referensi

- [1]. Castillo C, Mendoza M, Poblete B. Information credibility on twitter. In: Proceedings of the 20th international conference on World wide web [Internet]. New York, NY, USA: ACM; 2011 [cited 2023 Jun 17]. Available from: <http://dx.doi.org/10.1145/1963405.1963500>
- [2]. Dai Z, Yang Z, Yang Y, Carbonell J, Le Q, Salakhutdinov R. Transformer-XL: Attentive Language Models beyond a Fixed-Length Context. In: Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics [Internet].

- Stroudsburg, PA, USA: Association for Computational Linguistics; 2019 [cited 2023 Jun 17]. Available from: <http://dx.doi.org/10.18653/v1/p19-1285>
- [3]. Devlin J, Chang MW, Lee K, Toutanova K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding [Internet]. ACL Anthology. 2019 [cited 2023 Jun 17]. Available from: <https://aclanthology.org/N19-1423>
- [4]. Agnes. Studi Kasus Random Forest Machine Learning untuk Pemula Data. DQLab | Kursus Data Science Online Indonesia R Python [Internet]. 2022 Sep 21 [cited 2023 Jun 17]; Available from: <https://dqlab.id/studi-kasus-random-forest-machine-learning-untuk-pemula-data>
- [5]. Roihan A, Sunarya PA, Rafika AS. Pemanfaatan Machine Learning dalam Berbagai Bidang: Review paper. IJCIT (Indonesian Journal on Computer and Information Technology). 2020 May 1;5(1).
- [6]. Wada S, Takeda T, Okada K, Manabe S, Konishi S, Kamohara J, et al. Over-Sampling Effect in Pre-Training for Bidirectional Encoder Representations from Transformers (BERT) to Localize Medical BERT and Enhance Biomedical BERT (Preprint) [Internet]. JMIR Publications Inc.; 2022 Jul [cited 2023 Jun 17]. Available from: <http://dx.doi.org/10.2196/preprints.40992>
- [7]. Xu Y, Wei H, Lin M, Deng Y, Sheng K, Zhang M, et al. Transformers in computational visual media: A survey. Computational Visual Media. 2021 Oct 27;8(1):33–62.
- [8]. Badrinarayanan V, Kendall A, Cipolla R. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2017 Dec 1;39(12):2481–95.
- [9]. Sun C, Qiu X, Xu Y, Huang X. How to Fine-Tune BERT for Text Classification? In: Lecture Notes in Computer Science [Internet]. Cham: Springer International Publishing; 2019 [cited 2023 Jun 17]. p. 194–206. Available from: http://dx.doi.org/10.1007/978-3-030-32381-3_16
- [10]. Afifah L. Apa itu Confusion Matrix di Machine Learning? [Internet]. IlmudataPy. 2021 [cited 2023 Jun 17]. Available from: <https://ilmudatapy.com/apa-itu-confusion-matrix/>