**DETEKSI TEKS PROMOSI JUDI ONLINE MENGGUNAKAN AI DENGAN KOMBINASI
NLP DAN DEEP LEARNING****¹Samuel, ²Dedy Prasetya Kristiadi**^{1,2} Sistem Informasi, STMIK Kuwera

Jl. Kalideres Permai No.3C, RT.1/RW.14, Kalideres, Kec. Kalideres, Kota Jakarta Barat, Daerah Khusus Ibukota Jakarta, Indonesia 11830

Email: ¹imsamuel1605@gmail.com, ²dedyprasetya.kuwera@gmail.com**ABSTRAK**

Penyebaran konten ilegal, khususnya promosi judi online, semakin marak di internet seiring perkembangan teknologi digital. Untuk mengatasi ini, penelitian kami mengembangkan model kecerdasan buatan (*AI*) yang mampu mengklasifikasi dan mendeteksi teks promosi judi online secara otomatis. Metode yang digunakan melibatkan pemrosesan bahasa alami (*NLP*) untuk ekstraksi fitur teks, serta algoritma machine learning seperti *Naive Bayes*, *Support Vector Machine (SVM)*, dan model deep learning. Secara khusus, kami melakukan fine-tuning model *IndoBERT* dan Transformer pada tugas ini. Dataset yang digunakan terdiri dari kumpulan teks yang dikategorikan sebagai promosi judi dan non-judi, yang dikumpulkan dari media sosial, situs web, dan forum daring. Hasil evaluasi menunjukkan bahwa model Transformer, terutama setelah di-fine-tuning dengan *IndoBERT*, memberikan akurasi tertinggi dalam mengidentifikasi teks promosi judi online. Temuan ini menunjukkan bahwa pendekatan *AI* dengan kombinasi *NLP* dan *deep learning*, khususnya model berbasis *Transformer* yang di-fine-tuning pada konteks bahasa Indonesia seperti *IndoBERT*, sangat efektif untuk sistem penyaringan konten digital guna mendukung upaya pencegahan penyebaran konten ilegal di internet.

Kata Kunci: *Klasifikasi Teks, Judi Online, Kecerdasan Buatan, NLP, Deep Learning***1. PENDAHULUAN****1.1. Latar belakang**

Kebutuhan informasi yang terus berkembang di era digital saat ini berjalan seiring dengan penggunaan media sosial yang banyak ditemukan pada browser-browser yang ada. Oleh karena itu, ketergantungan terhadap media sosial juga semakin tinggi [1]. Sementara itu, dengan adanya ketergantungan tersebut tidak sedikit orang yang merasa sulit untuk melepaskan diri dari media sosial. Kemudahan yang didapat dalam berkomunikasi pada media sosial diantaranya adalah segera mendapat jawaban atas pertanyaan dari sebuah kejadian dan informasi yang update dari beberapa sumber informasi. Pada dasarnya, media sosial adalah sarana bagi individu untuk

saling berinteraksi, menciptakan, berbagi, dan bertukar informasi serta gagasan—baik melalui tulisan, gambar, maupun video—dalam sebuah jaringan dan komunitas virtual [2].

Promosi judi online seringkali disamarkan dalam bentuk teks yang seolah tidak mencurigakan, seperti undangan bermain game, diskon, tautan ke situs tertentu, atau penggunaan kata-kata yang disamarkan (misalnya "togel", "slot gacor", "spin gratis") [3]. Penyebarannya pun sangat cepat melalui media sosial, situs web, dan aplikasi pesan instan. Hal ini menyulitkan upaya penyaringan manual yang dilakukan oleh otoritas atau platform digital. Oleh karena itu, dibutuhkan pendekatan yang lebih cerdas dan otomatis untuk mengidentifikasi teks-teks semacam ini.

Salah satu solusi yang potensial adalah dengan menerapkan teknologi kecerdasan buatan (*Artificial Intelligence/AI*), khususnya teknik pemrosesan bahasa alami (*Natural Language Processing/NLP*)[4] dan pembelajaran mesin (*Machine Learning*)[5]. Berbagai penelitian sebelumnya telah membuktikan efektivitas model *AI* dalam klasifikasi teks, seperti untuk mendeteksi spam, ujaran kebencian, dan hoaks. Namun, penelitian khusus yang berfokus pada klasifikasi teks promosi judi online masih relatif terbatas, terutama dalam konteks bahasa Indonesia dan variasi gaya bahasa yang digunakan oleh pelaku promosi.

Penelitian ini bertujuan untuk mengembangkan dan mengevaluasi model *AI* yang mampu mengklasifikasi teks promosi judi online secara akurat. Penelitian mencakup tahapan pengumpulan dataset, pra-pemrosesan teks, ekstraksi fitur, pelatihan model klasifikasi, serta evaluasi kinerja model menggunakan metrik akurasi, presisi, *recall*, dan *F1-score*. Dengan adanya sistem klasifikasi ini, diharapkan dapat membantu pihak berwenang, pengelola platform, dan masyarakat umum dalam meminimalkan penyebaran konten ilegal dan melindungi ruang digital dari pengaruh negatif judi online.

2. LANDASAN TEORI

2.1. Kecerdasan Buatan (*Artificial Intelligence*)

Cabang ilmu komputer yang berfokus kepada pengembangan sistem komputer yang mampu melakukan tugas layaknya manusia. Tujuan utama dari kecerdasan buatan adalah membuat mesin yang dapat belajar, memahami, merencanakan dan beradaptasi sehingga dapat menyelesaikan tugas-tugas secara mandiri. [6]

2.2. Pembelajaran Mesin (*ML*)

Pendekatan dalam *AI* yang banyak digunakan untuk menggantikan atau menirukan perilaku manusia untuk menyelesaikan masalah atau melakukan otomatisasi. Sesuai namanya, *ML* mencoba menirukan bagaimana proses manusia atau makhluk cerdas belajar dan menggeneralisasi.1 Setidaknya ada dua aplikasi utama dalam *ML* yaitu, klasifikasi dan prediksi . [7]

2.3. Klasifikasi Teks

Klasifikasi teks adalah proses pemberian label kategori pada suatu dokumen teks berdasarkan isi atau ciri-ciri tertentu. Dalam konteks ini, klasifikasi dilakukan untuk membedakan antara

teks promosi judi dan teks non-judi. Proses klasifikasi umumnya terdiri dari beberapa tahap, yaitu pra-pemrosesan data, ekstraksi fitur (seperti *TF-IDF* atau *word embeddings*), pelatihan model, dan evaluasi kinerja model[13].

2.4. Deep Learning

Teknik dalam *NN* yang menggunakan teknik tertentu seperti *Restricted Boltzmann Machine (RBM)* untuk mempercepat proses pembelajaran dalam *NN* yang menggunakan lapis yang banyak atau lebih dari 7 lapis[9]. Dengan adanya *DL*, waktu yang dibutuhkan untuk *training* akan semakin sedikit karena masalah hilangnya gradien pada propagasi balik akan semakin rendah. [7]

2.5. Algoritma Naïve Bayes

Metode klasifikasi *Naïve Bayes* memanfaatkan prinsip probabilitas dan statistik yang dikembangkan oleh ilmuwan Inggris, *Thomas Bayes*. Pada intinya, metode ini menggunakan prinsip dasar *Teorema Bayes* untuk memprediksi kemungkinan kejadian di masa depan berdasarkan data dan pengalaman masa lalu [8]

2.6. Algoritma Support Vector Machine

Algoritma pembelajaran mesin yang dikenal sebagai *Support Vector Machine (SVM)* digunakan untuk menentukan kelas fitur data dan klasifikasi regresi Kemampuan *SVM* untuk menangani jumlah data yang sangat besar menjadikannya metode yang efektif dalam klasifikasi [8]

3. METODOLOGI PENELITIAN

Penelitian ini menggunakan pendekatan eksperimen komputasi untuk mengembangkan dan mengevaluasi model klasifikasi teks berbasis *Transformer* dalam mendeteksi komentar promosi [10] judi online. Metode yang digunakan mencakup proses mulai dari pengumpulan data hingga evaluasi performa model klasifikasi. Seluruh eksperimen dilakukan menggunakan bahasa pemrograman *Python* dengan pustaka *Hugging Face Transformers*, *scikit-learn*, dan *Datasets*[11][12].

3.1. Desain Penelitian

Penelitian ini bersifat kuantitatif dengan pendekatan *supervised learning*. Model dilatih untuk mengklasifikasikan teks ke dalam dua kelas: “*Gambling*” dan “*Non-Gambling*” berdasarkan komentar yang diambil dari platform *YouTube*. Model dikembangkan dengan arsitektur *BERT (Bidirectional*

Encoder Representations from Transformers) versi Bahasa Indonesia[12].

3.2. Pengumpulan Data

Data dikumpulkan dari komentar pada video YouTube menggunakan alat bantu seperti *YouTube Data API* dan web scraper pihak ketiga. Dataset akhir berisi ribuan entri komentar berbahasa Indonesia yang telah diberi label secara manual menjadi dua kelas:

- Label 1 – *Gambling*: komentar yang mengandung promosi judi online.
- Label 0 – *Non-Gambling*: komentar umum yang tidak mengandung unsur judi.

Dataset tersebut kemudian diunggah dan diproses menggunakan pustaka datasets dari *Hugging Face*. Untuk menjamin kualitas data, komentar kosong atau tidak valid disaring menggunakan fungsi khusus sebelum dilatih.

3.3. Tokenisasi dan Pra-pemrosesan

Model menggunakan *tokenizer* dari pretrained model *IndoBERT*. Setiap komentar diproses dengan:

- Tokenisasi: memecah teks menjadi *token BERT*.
- *Truncation dan Padding*: memastikan panjang *token* konsisten (maksimal 128 *token*).
- *Input Format Conversion*: data diubah menjadi format *tensor* (*input_ids*, *attention_mask*, dan *label*) untuk kebutuhan model.

3.4. Arsitektur dan Model

Model utama yang digunakan adalah *BertForSequenceClassification* dari pustaka *Hugging Face* dengan basis *IndoBERT* yang dilatih ulang (*fine-tuned*) menggunakan dataset spesifik ini. Model dikonfigurasi dengan dua *output* kelas (biner).

Pelatihan model dilakukan dengan parameter sebagai berikut:

- *Learning Rate*: 2e-5
- *Batch Size*: 16 (dengan *gradient accumulation step* 2)
- *Epochs*: 4
- *Loss Function*: *Cross Entropy*
- *Optimizer*: *AdamW*
- *FP16 Training*: Aktif untuk efisiensi memori
- *Evaluation Strategy*: per *epoch*
- *Early Stopping*: digunakan dengan *patience* = 2 untuk mencegah *overfitting*

Model dilatih menggunakan pustaka dari *Hugging Face* yang mengatur loop

pelatihan, evaluasi, dan penyimpanan model terbaik secara otomatis.

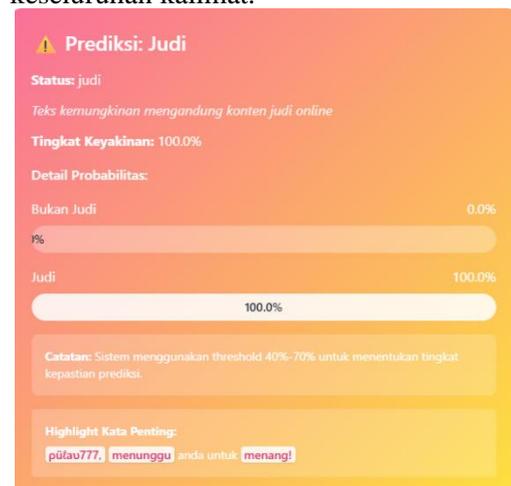
3.5. Evaluasi Model

Evaluasi performa model dilakukan pada data validasi menggunakan metrik:

- *Accuracy*
- *F1-Score*
- *Precision dan Recall* (melalui *classification_report*)

Hasil ini digunakan untuk menilai kemampuan model dalam membedakan teks yang mengandung promosi judi online dari komentar biasa.

Model yang tengah dikembangkan dalam penelitian ini dirancang untuk mengidentifikasi pola linguistik yang umum ditemukan pada konten promosi judi daring. Pendekatan yang digunakan adalah klasifikasi teks berbasis pembelajaran mesin, di mana model tidak hanya mempertimbangkan keberadaan kata kunci secara individual, tetapi juga memerhatikan konteks semantik dari keseluruhan kalimat.



Gambar 1. Evaluasi Model

Sebagaimana ditunjukkan pada Gambar 1, model mampu memberikan prediksi dengan tingkat keyakinan mencapai 100% terhadap sebuah komentar yang diduga mengandung unsur judi. Hal ini didasarkan pada kemunculan sekumpulan kata yang secara statistik memiliki keterkaitan kuat dengan praktik promosi perjudian, seperti:

- “püäu777” — representasi tipikal dari domain atau nama situs judi yang sengaja dimodifikasi untuk menghindari sistem penyaringan otomatis.
- “menunggu” — kata yang digunakan dalam strategi persuasi untuk

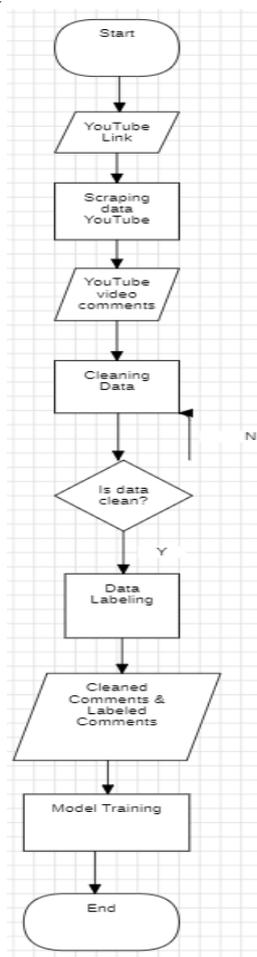
membangun urgensi atau rasa penasaran.

- “menang!” — kata yang secara langsung memicu asosiasi terhadap keuntungan atau hadiah, yang merupakan karakteristik umum dari promosi perjudian.

Model memperoleh pengetahuan tersebut melalui proses pelatihan berbasis data yang telah dilabeli sebelumnya. Selama proses ini, model belajar mengenali kombinasi kata dan frasa yang secara konsisten muncul dalam konteks yang mengindikasikan aktivitas perjudian.

Dalam pengembangan awal ini, sistem menggunakan ambang batas probabilitas (threshold) antara 40% hingga 70% sebagai zona ketidakpastian untuk menilai kepercayaan terhadap prediksi. Apabila nilai probabilitas melebihi ambang tersebut, seperti pada kasus ini (100%), maka sistem mengklasifikasikan komentar sebagai mengandung konten judi dengan tingkat kepastian yang tinggi.

3.6. Flowchart



Gambar 2. Flowchart

Flowchart pada Gambar 2 menjelaskan tahapan-tahapan sistematis dalam proses pengumpulan, pembersihan, pelabelan, hingga pelatihan model menggunakan data komentar dari YouTube. Adapun penjelasan dari tiap tahapan adalah sebagai berikut:

3.6.1. Start

Tahap awal dimulainya proses secara keseluruhan.

3.6.2. YouTube Link

Pengguna memasukkan tautan video YouTube yang menjadi sumber data komentar.

3.6.3. Scraping Data YouTube

Pada tahap ini, dilakukan proses pengambilan komentar dari video yang telah ditentukan menggunakan teknik *scraping*, baik melalui *API YouTube* maupun metode lain yang relevan.

3.6.4. YouTube video comments

Komentar yang telah berhasil dikumpulkan disimpan sebagai data mentah untuk kemudian diproses lebih lanjut.

3.6.5. Cleaning Data

Data komentar dibersihkan dari berbagai elemen yang tidak dibutuhkan seperti karakter spesial, emoji, *HTML tag*, dan elemen lain yang dapat mengganggu proses analisis.

3.6.6. Is data clean?

Dilakukan pengecekan terhadap data hasil *cleaning* untuk memastikan data sudah layak untuk tahap pelabelan. Jika ditemukan data yang belum bersih, maka proses kembali ke tahap *Cleaning Data*.

3.6.7. Data Labeling

Data yang telah bersih kemudian diberi label sesuai dengan kategori yang ditentukan, seperti sentimen positif, negatif, netral, atau kategori tematik lainnya seperti indikasi unsur judi, spam, dan sebagainya.

3.6.8. Cleaned Comments & Labeled Comments

Data yang telah melalui proses pembersihan dan pelabelan ini disiapkan sebagai dataset final yang akan digunakan untuk pelatihan model.

3.6.9. Model Training

Dataset final digunakan untuk melatih model pembelajaran mesin (*machine learning*) agar dapat melakukan klasifikasi atau prediksi terhadap komentar baru secara otomatis.

3.6.10. End

Proses berakhir setelah model berhasil dilatih.

4. Hasil dan Pembahasan

Hasil evaluasi menunjukkan performa model klasifikasi teks dalam mendeteksi konten perjudian dan non-perjudian pada data berbahasa Indonesia. Evaluasi dilakukan menggunakan metrik *precision*, *recall*, *f1-score*, dan akurasi untuk menilai sejauh mana model mampu membedakan kedua kategori secara efektif. Pembahasan difokuskan pada interpretasi metrik tersebut serta implikasinya terhadap penerapan model dalam skenario dunia nyata.

	precision	recall	f1-score	support
Non-Gambling	0.99	0.93	0.96	4456
Gambling	0.96	0.99	0.97	6864
accuracy			0.97	11320
macro avg	0.97	0.96	0.97	11320
weighted avg	0.97	0.97	0.97	11320

Gambar 3. *Classification Report* setelah *training*

Model klasifikasi yang dikembangkan menunjukkan performa yang sangat baik dalam membedakan antara konten perjudian dan non-perjudian pada teks berbahasa Indonesia. Berdasarkan hasil evaluasi pada data uji, diperoleh nilai akurasi sebesar 97%, yang mencerminkan proporsi prediksi model yang benar terhadap seluruh data uji.

Secara lebih rinci, kelas *Gambling* memiliki nilai *recall* sebesar 0.99, yang menunjukkan bahwa hampir seluruh data yang termasuk dalam kategori ini berhasil dikenali oleh model. Sementara itu, kelas *Non-Gambling* memperoleh *precision* sebesar 0.99, yang mengindikasikan bahwa sebagian besar prediksi model untuk kelas ini sesuai dengan label sebenarnya. Nilai *f1-score* pada kedua kelas berturut-turut adalah 0.97 untuk *Gambling* dan 0.96 untuk *Non-Gambling*, yang menunjukkan keseimbangan antara *precision* dan *recall* dalam pengklasifikasian kedua kelas tersebut.

Nilai rata-rata makro (*macro average*) dan berbobot (*weighted average*) masing-masing sebesar 0.97 untuk *precision* dan *f1-score*, serta 0.96–0.97 untuk *recall*. Hal ini mengindikasikan bahwa model mampu mempertahankan kinerja yang konsisten pada kedua kelas, meskipun jumlah data antar kelas tidak sepenuhnya seimbang.

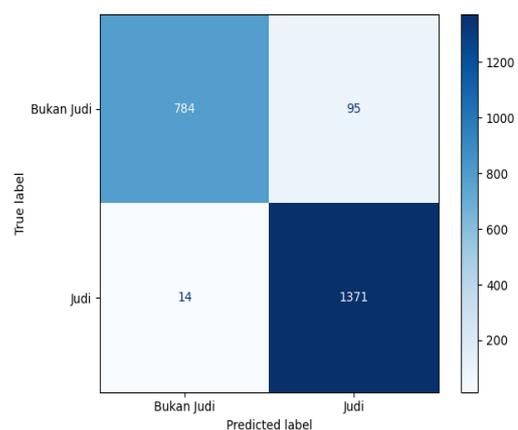
Secara keseluruhan, hasil ini menunjukkan bahwa model yang digunakan cukup andal untuk diterapkan dalam tugas-tugas klasifikasi teks yang berkaitan dengan deteksi konten perjudian. Tingginya nilai *recall* pada kelas *Gambling* menjadi aspek penting dalam konteks deteksi konten bermasalah, di mana ketepatan dalam menangkap kasus positif sangat krusial untuk menghindari kesalahan klasifikasi yang berpotensi berdampak pada sistem moderasi atau pengambilan keputusan.

4.1. Hasil Performance

Data uji yang digunakan merupakan sampel teks yang diambil dari video *YouTube*, di mana data ini belum pernah dilibatkan dalam proses pelatihan (*training*), sehingga dapat dianggap sebagai data murni untuk evaluasi model. Evaluasi dilakukan terhadap dua model klasifikasi teks, yaitu *Naïve Bayes* dan *Support Vector Machine* (SVM).

Total data uji yang digunakan sebanyak 2.264 sampel, yang terdiri dari:

- 979 sampel kategori "Bukan Judi"
- 1.285 sampel kategori "Judi"



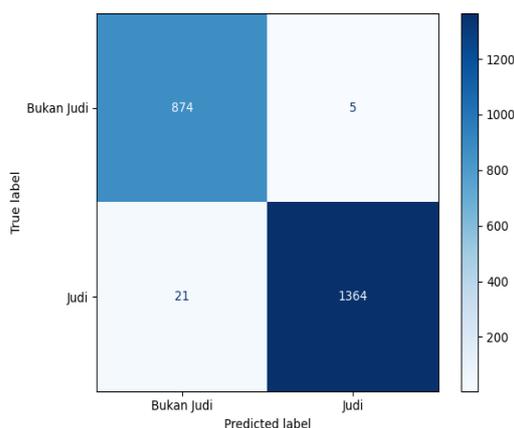
Gambar 4. *Confusion Matrix Naïve Bayes*

Gambar 4 menunjukkan confusion matrix hasil prediksi model *Naïve Bayes*. Model ini menghasilkan:

- *True Positive (TP)*: 1.371 data "Judi" berhasil diklasifikasikan dengan benar.

- *True Negative (TN)*: 784 data "Bukan Judi" berhasil diklasifikasikan dengan benar.
- *False Positive (FP)*: 95 data "Bukan Judi" salah diklasifikasikan sebagai "Judi".
- *False Negative (FN)*: 14 data "Judi" salah diklasifikasikan sebagai "Bukan Judi".

Model Naïve Bayes mampu mengklasifikasikan sebagian besar data dengan cukup baik, namun masih terdapat 95 sampel komentar non-judi yang salah diklasifikasikan sebagai komentar judi (false positive). Di sisi lain, hanya 14 komentar judi yang tidak terdeteksi sebagai judi (false negative). Hal ini menunjukkan bahwa model ini cenderung lebih sensitif terhadap komentar yang mengandung indikasi judi, namun kurang presisi dalam membedakan komentar yang sebenarnya tidak mengandung unsur judi.



Gambar 5. *Confusion Matrix SVM*

Sebagai pembandingan, model SVM menunjukkan performa yang lebih baik, dengan hasil:

- *True Positive (TP)*: 1.364
- *True Negative (TN)*: 874
- *False Positive (FP)*: 5
- *False Negative (FN)*: 21

Dengan hanya 5 kesalahan pada prediksi "Bukan Judi" dan akurasi total sekitar 98.3%, SVM memperlihatkan keseimbangan yang sangat baik antara presisi dan *recall*. Model ini lebih andal untuk diterapkan dalam sistem deteksi otomatis, khususnya untuk konteks yang sensitif terhadap kesalahan klasifikasi,

seperti moderasi konten publik atau filtering pada media sosial.



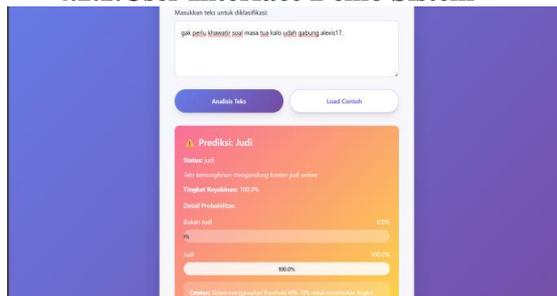
Gambar 6. *Word Cloud*

Word cloud di atas menampilkan representasi visual dari kata-kata yang paling sering muncul dalam komentar yang terklasifikasi sebagai judi. Terlihat bahwa kata "alexis17" mendominasi secara signifikan, mengindikasikan bahwa nama tersebut sangat sering disebut dalam konteks aktivitas perjudian. Selain itu, kata-kata seperti "main", "fitur", "event", "join", dan "bonus" juga muncul cukup besar, menandakan adanya promosi atau ajakan untuk bergabung ke suatu platform. Kata ganti orang seperti "gue", "kamu", dan "semua" juga menunjukkan gaya bahasa informal yang lazim di media sosial. *Word cloud* ini membantu memperjelas bahwa komentar judi tidak hanya menyebut nama platform, tetapi juga menyertakan ajakan, pengalaman bermain, serta istilah khas dunia perjudian daring.

4.2. Perancangan Demo Sistem

Untuk menunjukkan implementasi dan kinerja dari model klasifikasi yang telah dikembangkan, dilakukan perancangan dan pengujian melalui demo sistem. Sistem ini bertujuan untuk memperlihatkan bagaimana model bekerja dalam mengidentifikasi teks yang mengandung unsur perjudian secara otomatis. Demo sistem dikembangkan sebagai alat bantu visual sekaligus validasi terhadap hasil pelatihan model yang telah dilakukan sebelumnya.

4.2.1. User Interface Demo Sistem



Gambar 7. Gambar *User Interface* Demo Sistem

5. Penutup

Tujuan Penelitian adalah ini untuk mengembangkan model klasifikasi teks berbasis kecerdasan buatan yang dapat mendeteksi komentar-komentar yang mengandung unsur promosi judi online, khususnya dalam bahasa Indonesia. Dengan memanfaatkan pendekatan pemrosesan bahasa alami (NLP) dan algoritma machine learning, termasuk *Naïve Bayes*, *Support Vector Machine (SVM)*, serta model *deep learning* berbasis Transformer (*IndoBERT*), sistem yang dibangun menunjukkan performa klasifikasi yang tinggi. Model *Transformer* yang di-*fine-tune* dengan *IndoBERT* mencatat akurasi tertinggi, mencapai 97% dengan nilai *precision* dan *recall* yang seimbang. Sementara itu, SVM juga menunjukkan performa sangat baik dengan tingkat kesalahan yang minim. *Word cloud* dan evaluasi metrik lainnya semakin memperkuat bahwa pendekatan ini mampu mengenali pola linguistik khas yang sering digunakan dalam promosi judi. Dengan demikian, hasil penelitian ini menunjukkan bahwa teknologi AI dapat dimanfaatkan secara efektif sebagai alat bantu otomatis dalam proses moderasi konten digital, khususnya untuk memerangi penyebaran konten perjudian di platform daring seperti *YouTube*.

DAFTAR PUSTAKA

- [1] Tryas Amanda Putri, Juniar Gasa Nova, Tesa Amilia Putri,dkk,2025. Tren Penelitian Dampak Kecanduan Media Sosial Bagi Remaja, *Jurnal Ilmu Kesehatan Umum, Psikolog, Keperawatan dan Kebidanan*, Volume.3, Nomor.2 Juni 2025
- [2] Dharmawan, S., Viny,), Mawardi, C., Novario,), & Perdana, J. (2023). *Jurnal Ilmu Komputer dan Sistem Informasi Klasifikasi Ujaran Kebencian Menggunakan Metode FeedForward Neural Network (IndoBERT)*.
- [3] Jalin, 2025. Indonesia Darurat Judi Online. <https://www.jalin.co.id/id-id/berita/blog/indonesia-darurat-judi-online-simak-penyebab-dan-solusi-yang-dapat-diterapkan>.
- [4] Raditya Danar,dkk, 2024. Dasar-dasar Natural Language Processing(BLP),Minhaj Pustaka,2024.
- [5] Agung Wijoyo,dkk,2024. Pembelajaran Machine Learning, OKTAL : Jurnal Ilmu Komputer dan Science Volume 3, No. 2, Februari 2024
- [6] Karyadi, B. (2023). *PEMANFAATAN KECERDASAN BUATAN DALAM MENDUKUNG PEMBELAJARAN MANDIRI*. 8(2), 253–258. <https://doi.org/10.32832/educate.v8i02.14843>
- [7] Ahmad, A. (2017). *Mengenal Artificial Intelligence, Machine Learning, Neural Network, dan Deep Learning*. www.teknoindonesia.com
- [8] Maulana, A., & Yuliana, A. (2024). ANALISIS SENTIMEN OPINI PUBLIK TERKAIT JUDI ONLINE PADA PENGGUNA APLIKASI X MENGGUNAKAN ALGORITMA NAÏVE BAYES DAN SUPPORT VECTOR MECHINE. *Jurnal Informatika dan Teknik Elektro Terapan*, 12(3S1). <https://doi.org/10.23960/jitet.v12i3S1.5187>
- [9] Hendi Hidayat1, Freza Riana,2023,Identifikasi Benih Jahe menggunakan Convolutional Neural Network, *Jurnal Infotech*,Vol. 9 No. 1, pp. 287-298.
- [10] Muhammad Basil Musyaffa Amin, Gibran Hakim, dkk, 2019. DETEKSI SPAM BERBAHASA INDONESIA BERBASIS TEKS MENGGUNAKAN MODEL BERT ,*Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK)* Vol. 11, No. 6, Desember 2024, hlm. 1291-1302.
- [11] Chaumond, J., Delangue, C., & Wolf, T. (2016). *huggingface (Hugging Face)*. <https://huggingface.co/huggingface>
- [12] Devlin, J., Chang, M.-W., Lee, K., Google, K. T., & Language, A. I.,2018. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. <http://arxiv.org/abs/1810.04805>
- [13] Alwehaibi, A., Bikdash, M., Albogmi, M., & Roy, K. (2021). A study of the performance of embedding methods for Arabic short-text sentiment analysis using deep learning approaches. *Journal of King Saud University-Computer and Information Sciences*