

BAB I

PENDAHULUAN

1.1 Latar Belakang

Kasus perjudian di Indonesia merupakan permasalahan sosial yang telah berlangsung lama dan semakin kompleks seiring dengan perkembangan teknologi digital. Saat ini, praktik perjudian telah bertransformasi menjadi judi online, yang penyebarannya berlangsung secara masif melalui media sosial. Pada akhir tahun 2023, Pusat Pelaporan dan Analisis Transaksi Keuangan (PPATK) mencatat total transaksi terkait judi online mencapai Rp327 triliun, dengan 2,37 juta pengguna aktif terlibat di dalamnya [1].

Kemudahan akses terhadap media sosial mendorong penyebaran konten promosi judi online secara luas. Studi menunjukkan bahwa 82% pengguna internet pernah melihat iklan judi online, dengan platform seperti Instagram, Facebook, dan YouTube menjadi saluran utama penyebarannya. Konten tersebut umumnya berbentuk teks yang sulit dideteksi secara otomatis, seperti penggunaan kata tidak baku, simbol pengganti huruf, atau ejaan yang dimanipulasi [1][2].

Dalam merespons fenomena ini, pemerintah Indonesia mulai mengadopsi teknologi berbasis Artificial Intelligence (AI) dan machine learning sebagai bagian dari strategi deteksi dan pencegahan dini [2][19][20]. Teknologi ini dinilai mampu mempercepat identifikasi konten ilegal dalam skala besar, serta memperbarui metode deteksi seiring dengan evolusi strategi pelaku. Namun demikian, penegakan hukum terhadap judi online di Indonesia masih menghadapi berbagai tantangan struktural dan teknis, seperti kesenjangan regulasi, keterbatasan dalam pelacakan identitas digital, serta belum optimalnya sistem deteksi otomatis [27].

Penelitian oleh Perdana et al. (2024) menunjukkan bahwa promosi judi online dapat dideteksi menggunakan pendekatan machine learning seperti Random Forest, Logistic Regression, dan Convolutional Neural Networks (CNN) dengan akurasi tinggi pada dataset Twitter. Namun, penelitian tersebut juga merekomendasikan

eksplorasi model berbasis deep learning untuk menangani variasi konteks dan manipulasi bahasa yang lebih kompleks [11].

Tantangan utama dalam deteksi teks promosi judi online memang terletak pada variasi struktur bahasa, bentuk kalimat informal, dan manipulasi ejaan. Hal ini sejalan dengan studi oleh Khafid et al. (2024) dalam Jurnal JTIK, yang menunjukkan bahwa meskipun model BERT menghasilkan akurasi tinggi (94%) dan F1-score di atas 0,92 pada teks standar, performanya menurun signifikan saat dihadapkan dengan teks berisi ejaan tidak baku, sebagaimana umum ditemukan dalam komentar promosi judi online [21].

Studi oleh Meftouh et al. (2022) dalam ACL Anthology turut memperkuat temuan ini. Penelitian tersebut menyatakan bahwa model transformer seperti BERT mengalami degradasi performa hingga 25–30% ketika digunakan untuk memproses teks dari media sosial yang mengandung slang, simbolisasi, serta pemenggalan kata—terutama pada bahasa non-Inggris seperti Bahasa Indonesia yang memiliki ragam bentuk penulisan informal [30]. Oleh karena itu, pendekatan Natural Language Processing (NLP) seperti BERT tetap membutuhkan dukungan metode pelengkap yang toleran terhadap kesalahan ejaan dan variasi penulisan.

IndoBERT, sebagai varian lokal dari BERT, telah terbukti lebih akurat dalam memahami konteks kalimat dalam Bahasa Indonesia [3][22]. Akan tetapi, IndoBERT tetap memiliki keterbatasan dalam menangani bentuk teks yang dimanipulasi, seperti penggunaan angka, simbol, atau gabungan karakter sebagai pengganti huruf. Untuk mengatasi hal ini, Fuzzy Matching dapat digunakan sebagai filter awal yang mengenali kemiripan kata berbasis string similarity [4][23]. Teknik ini terbukti efektif dalam mengidentifikasi pola manipulatif yang sering digunakan dalam teks promosi judi online [24].

Penelitian oleh Arslan dan Cruz (2023) juga menyatakan pentingnya untuk menguji beragam skenario klasifikasi dan strategi pelabelan secara sistematis. Ini agar model dapat menangani kompleksitas struktur label dan distribusi data yang tidak seimbang. Melalui pendekatan multimetode, model dapat dilatih untuk

menangani representasi teks yang kompleks dan lebih adaptif terhadap kondisi dunia nyata, seperti komentar promosi judi online yang tidak terstruktur [31].

Berdasarkan latar belakang dan penelitian terdahulu yang telah dipaparkan, penelitian ini akan menggabungkan algoritma IndoBERT dan Fuzzy Matching ke dalam sebuah pipeline hybrid, serta melakukan evaluasi terhadap efektivitas gabungan keduanya dalam mendeteksi teks judi online—terutama dalam kondisi manipulasi bahasa dan variasi penulisan yang umum ditemukan pada konten media sosial.

1.2 Identifikasi Masalah

Berdasarkan latar belakang yang sudah tertulis, didapati identifikasi masalah sebagai berikut:

- Bagaimana performa algoritma IndoBERT dalam mendeteksi teks judi online?
- Bagaimana performa gabungan antara IndoBERT dan *Fuzzy Matching* dalam mendeteksi judi online?

1.3 Batasan Penelitian

Penelitian ini memiliki beberapa batasan, antara lain:

- Penelitian ini tidak membuat *user interface*.
- Dataset yang digunakan adalah dataset publik.
- Penelitian dilakukan pada platform *Jupyterlab*.

1.4 Tujuan Penelitian

Penelitian ini bertujuan untuk menerapkan, menganalisis, dan mengevaluasi performa sistem hybrid berbasis IndoBERT dan Fuzzy Matching dalam mendeteksi teks yang mengandung unsur judi online, serta mendapat akurasi terbaik meskipun terdapat variasi konteks penulisan.

1.5 Manfaat Penelitian

Penelitian ini memiliki beberapa manfaat, antara lain:

- Memberikan solusi teknologi yang dapat meningkatkan efektivitas dalam mendeteksi dan mengurangi penawaran judi online dalam bentuk teks pada media sosial.
- Memperluas dataset komentar yang mengandung kata atau kalimat judi online yang dapat diakses secara publik.
- Dapat digunakan sebagai sumber referensi akademik dan teknis bagi peneliti lain dalam mengembangkan teknologi Machine Learning.

1.6 Sistematika Penulisan

Berikut adalah sistematika penulisan yang digunakan dalam penelitian ini:

BAB I PENDAHULUAN

Pembahasan mengenai lingkup latar belakang, identifikasi masalah, batasan penelitian, tujuan penelitian, dan manfaat penelitian.

BAB II TINJAUAN PUSTAKA

Pada bab ini, terdapat uraian mengenai bahan-bahan kajian teori seperti penelitian terdahulu tentang penggunaan IndoBERT, *Fuzzy Matching*, dan teknologi AI dalam menangani judi online. Bab ini berguna sebagai landasan bagi peneliti dalam merumuskan solusi terhadap masalah yang diteliti.

BAB III ANALISIS DAN PERANCANGAN SISTEM

Pada bab ini membahas mengenai rancangan jalannya penelitian beserta alur penelitian yang terdiri dari berbagai skenario atau case dan juga menggunakan penjelasan flowchart.

BAB IV IMPLEMENTASI DAN PENGUJIAN SISTEM

Pada Bab ini membahas tentang implementasi sistem, termasuk proses pembangunan, pengujian, dan penyesuaian sistem dengan maksud dan tujuan penelitian.

BAB V SIMPULAN DAN SARAN

Pada bab ini membahas mengenai kesimpulan akhir dan saran yang didapat dari penelitian yang dilakukan



UMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA