

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Berita *online* adalah informasi yang dapat dijangkau melalui internet. Ini adalah tipe berita yang baru. Di Indonesia, situs berita *online* Detik.com adalah hasil produksi dari PT Agrakom, yang merupakan tempat pertama dalam kategori berita *online*. Detik.com menyediakan informasi mengenai kejadian terbaru. Selain itu, Detik.com juga menyajikan berbagai informasi lainnya dalam berbagai bidang dan lebih banyak lagi. Seiring berjalannya waktu, Detik.com terus beradaptasi dengan kebutuhan yang berkembang dari setiap pembacanya (Andriani & Putri, 2020).

Pada era modern ini, penggunaan judi *online* semakin meluas. Di Kantor Direktorat Jenderal Bea dan Cukai di Jakarta Timur pada hari Kamis 14 November, dia menyatakan bahwa dari data yang dikumpulkan oleh intelijen ekonomi, sebanyak 8,8 juta pemain akan berpartisipasi dalam permainan judi online pada tahun 2024, dengan 80 persen dari pemain tersebut berasal dari masyarakat bawah dan fokus pada anak-anak muda (CNN Indonesia-Budi Gunawan, 2024).

Penggunaan perjudian *online* sedang meningkat bukan cuma di Indonesia namun juga di banyak negara di seluruh dunia. Pada kondisi tersebut, media menjadi sarana vital untuk para pencinta judi *online* guna menyampaikan pendapat dan komentarnya mengenai pengalamannya terkait aktivitas tersebut. Perkembangan pesat di sektor teknologi informasi memberikan efek besar pada banyak bidang kehidupan. Satu diantara dampak paling signifikan dari perkembangan pesat di sektor informatika ialah keahlian dalam menghimpun dan menganalisis data berskala besar dan cepat daripada sebelumnya. Pada hal analisis sentimen untuk pengguna perjudian *online*, ini bermakna mampu mengimpun dan mengolah data dari pesan-pesan di detik.com secara efisien. Metode NLP bisa dijalankan guna menjalankan analisis teks otomatis dan memperoleh pemahaman berharga tentang sikap dan preferensi pengguna terhadap judi *online*. Lewat analisis sentimen, bisa dipahami bahwasanya efek psikologis dan sosial dari judi *online*. Metode analisis yang dipakai pada riset berikut ialah *Natural Language Processing*,

yakni suatu teknologi yang memungkinkan komputer memahami manusia secara alami dan mengolah informasi yang dikandungnya (Julianti & Suarna, 2024).

Algoritma yang dipakai guna menganalisis pandangan pada permasalahan judi *online* ialah *Decision Tree* yang merupakan suatu pendekatan yang sangat dikenal dan praktis dalam pembelajaran mesin untuk menangani masalah klasifikasi, Selain karena proses pengembangannya tergolong cepat, hasil dari model yang diciptakan juga simpel untuk dipahami (Anam & Firdaus, 2021). kinerja model yang optimal sering kali bergantung pada penyetelan *hyperparameter* yang tepat. *GridSearchCV* adalah sebuah teknik pengoptimalan *hyperparameter* yang secara sistematis menelusuri kisi *hyperparameter* yang telah ditetapkan sebelumnya untuk mengidentifikasi kombinasi parameter terbaik bagi model. *GridSearchCV* telah diadopsi secara luas dalam penelitian pembelajaran mesin karena kemampuannya untuk menyempurnakan model dan meningkatkan akurasi prediktifnya (Rasheed & Kumar, 2024).

Penelitian ini dilatarbelakangi oleh terbatasnya kajian yang menganalisis sentimen teks berita mengenai judi *online* dengan pendekatan *machine learning*, khususnya algoritma *Decision Tree*, dan belum dibandingkan dengan algoritma lain yang lebih canggih seperti *Naïve Bayes* atau *Random Forest*, yang memiliki akurasi sekitar 75-78%. Penelitian ini mengisi celah tersebut dengan menerapkan NLP yang lebih akurat, mengeksplorasi metode klasifikasi *Decision Tree* yang lebih optimal, serta menganalisis keterkaitan persepsi publik dengan dampak sosial judi *online*.

## 1.2 Perumusan Masalah

Bersumber pemaparan latar belakang, rumusan permasalahan yang akan dijawab pada riset berikut ialah:

1. Bagaimana menerapkan *Natural Language Processing* (NLP) untuk menganalisis sentimen terkait judi *online* di Detik.com?
2. Bagaimana performa algoritma *Decision Tree* dalam melakukan analisis sentimen terhadap berita judi *online* di Detik.com tanpa menggunakan teknik optimasi?

3. Pendekatan optimasi seperti apa yang efektif dalam meningkatkan kinerja algoritma *Decision Tree* untuk analisis sentimen pada berita judi *online* di Detik.com?

### 1.3 Batasan Masalah

Untuk menjaga fokus dan keterbatasan waktu serta sumber daya, riset berikut terbatas pada sejumlah hal diantaranya:

1. Data yang dianalisis pada riset berikut terbatas pada berita yang terdapat di detik.com terkait dengan artikel tentang judi *online*.
2. Analisis sentimen hanya dilakukan memakai metode *Term Frequency-Inverse Document Frequency* (TF-IDF) sebagai representasi teks dan algoritma *Decision Tree* sebagai model klasifikasi.
3. Penelitian ini hanya menggolongkan sentimen ke dalam tiga kategori, yakni negatif, positif, dan netral.
4. Data yang dianalisis diambil dari periode waktu tertentu dan bersumber dari satu platform berita *online*, yaitu detik.com.
5. Validasi performa model akan menggunakan *matrix* evaluasi berupa *accuracy*, *presisi*, dan *recall* sebagai pengukuran utama untuk menilai kualitas hasil klasifikasi sentimen.

### 1.4 Tujuan

Bagian ini menjelaskan tujuan umum dan khusus yang berorientasi keilmuan yang akan dicapai pada riset berikut meliputi:

- a. Menerapkan *Natural Language Processing* (NLP) dalam analisis sentimen terkait judi *online* di detik.com
- b. Mengevaluasi performa algoritma *Decision tree* dalam analisis sentimen terkait judi *online* di detik.com tanpa menggunakan optimasi berupa nilai *accuracy*, *presisi* dan *recall*.
- c. Meningkatkan akurasi algoritma *Decision Tree* dalam analisis sentimen terkait judi *online* di detik.com melalui optimasi *hyperparameter* menggunakan *GridSearchCV* berupa nilai *accuracy*, *presisi* dan *recall*.

## 1.5 Manfaat

Manfaat riset berikut yang diharapkan penulis meliputi:

### 1.5.1 Bagi Penulis

Menerapkan ilmu dalam melakukan analisis sentimen menggunakan *Natural Language Processing* (NLP) dan algoritma *Decision Tree*.

### 1.5.2 Bagi Pembaca

- a. Pembaca dapat memperoleh informasi tentang bagaimana masyarakat menanggapi berita mengenai judi *online*, baik dalam sentimen positif, negatif, maupun netral, sehingga dapat memahami tren persepsi publik secara lebih mendalam.
- b. Informasi hasil analisis sentimen berikut bisa dipakai menjadi acuan atau pedoman untuk penelitian atau penentuan keputusan yang berhubungan dengan permasalahan judi *online*.

### 1.5.3 Bagi Universitas

Memahami tolak ukur keahlian mahasiswa dalam menguasai materi yang didapat selama perkuliahan dan dapat menambah referensi literatur untuk riset selanjutnya.

## 1.6 Metodologi Penelitian

Alur riset yang dijalankan akan memberikan struktur sistematis dan membantu merencanakan, melaksanakan, dan menganalisis penelitian secara terstruktur.

### 1.6.1 Identifikasi Masalah

Penelitian ini mengidentifikasi bahwa judi online menjadi isu yang sering diberitakan di media seperti Detik.com. Namun, belum banyak penelitian yang secara khusus menganalisis sentimen berita terkait isu ini. Selain itu, masih diperlukan eksplorasi efektivitas algoritma *Decision Tree* dibandingkan metode lain dalam analisis sentimen berita.

### **1.6.2 Studi Literatur**

Studi literatur dijalankan dengan cara mengumpulkan dan mempelajari artikel, buku, dan jurnal ilmiah yang berkaitan dengan judul seperti analisis sentimen, NLP (Natural Language Processing), Decision Tree, dan hyperparameter menggunakan GridSearchCV, serta pengaplikasian bahasa pemrograman python dalam analisis sentimen. Riset berikut juga menggunakan jurnal, e-book dan website resmi sebagai referensi pada penulisan skripsi ini sehingga membantu dalam menyelesaikan penelitian ini.

### **1.6.3 Perancangan Sistem**

Pada tahap ini, penelitian mulai masuk ke aspek teknis dalam membangun sistem analisis sentimen terhadap pemberitaan judi *online*. Langkah-langkah yang dilakukan meliputi akuisisi data, *preprocessing data*, pelabelan sentimen, *feature extraction*, pemilihan model, optimasi model, perhitungan *confusion matrix* dan evaluasi model.

### **1.6.4 Pengujian**

Model *Decision Tree* yang telah dibangun diuji menggunakan dataset yang telah dikumpulkan untuk memastikan kinerjanya dengan menambahkan optimasi untuk meningkatkan akurasi dalam menganalisis sentimen berita terkait judi *online*.

### **1.6.5 Laporan**

Tahap ini merupakan proses akhir dari penelitian, di mana seluruh hasil, metode, serta analisis yang telah dilakukan didokumentasikan dalam bentuk laporan tugas akhir. Laporan ini berfungsi sebagai bukti ilmiah dari penelitian yang telah dilakukan serta sebagai referensi bagi penelitian selanjutnya.

## **1.7 Sistematika Penulisan**

Penelitian ini disusun secara sistematis ke dalam beberapa bab yang saling berkaitan satu sama lain untuk memberikan alur yang runtut. Sistematika ini juga

dimaksudkan untuk mempermudah pembaca memahami setiap langkah proses penelitian sebagai berikut.

#### **BAB I. PENDAHULUAN**

Pada bab ini tercakup penjelasan secara singkat mengenai latar belakang, rumusan masalah, batasan masalah, tujuan, manfaat, metode penelitian dan sistematika pembahasan.

#### **BAB II. TINJUAN PUSTAKA**

Dalam bab ini dibahas mengenai dasar-dasar teori yang menjadi landasan penelitian serta permasalahan yang pernah diteliti sebelumnya, beserta rencana yang akan dilakukan dalam penelitian selanjutnya. Penelitian ini berfokus pada topik analisis sentimen terhadap isu judi *online* dengan membandingkan performa algoritma *Decision Tree* dengan beberapa algoritma klasifikasi lainnya guna memperoleh hasil klasifikasi yang paling optimal.

#### **BAB III. METODE PENELITIAN**

Dalam bab ini membahas tentang metodologi beserta langkah-langkah yang digunakan peneliti yaitu alur penelitian, metode implementasi dan skenario pengujian. Untuk analisis data menggunakan algoritma *decision tree* dan *GridSearchCV* untuk optimasi.

#### **BAB IV. HASIL DAN PEMBAHASAN**

Dalam bab ini dibahas hasil penelitian yang berupa analisis sentimen dengan topik judi online yang dilakukan terhadap data dari situs Detik.com, menggunakan metodologi yang telah dijelaskan pada bab sebelumnya.

#### **BAB V. PENUTUP**

Dalam bab ini berisi kesimpulan dari hasil penelitian yang diperoleh berdasarkan rumusan masalah yang telah dirumuskan, untuk menjawab tujuan penelitian. Selain itu, bab ini juga terdapat saran

yang disusun berdasarkan manfaat penelitian, sebagai rekomendasi dan arahan untuk pengembangan penelitian selanjutnya