

BAB I

PENDAHULUAN

1.1 Latar Belakang

Kemajuan zaman menuntut penyajian informasi yang cepat, akurat, dan terpercaya. Pesatnya perkembangan teknologi informasi di masyarakat turut mendukung kemudahan dalam mengakses berbagai jenis informasi. Salah satu bentuk penyampaian informasi yang banyak diminati adalah berita. Berita merupakan salah satu bentuk komunikasi yang menyampaikan fakta atau opini, baik secara lisan maupun tertulis. Media penyajian berita pun mengalami transformasi, dari media cetak menjadi media digital. Di tengah kemajuan teknologi tersebut, berita olahraga menjadi salah satu topik yang banyak dicari dan diakses oleh masyarakat luas. Besarnya volume berita yang tersebar membutuhkan sistem klasifikasi otomatis untuk mengelompokkan teks berita olahraga secara efisien [1].

Klasifikasi teks adalah suatu proses penyuntingan sebagian besar informasi penting dari sumber teks. Klasifikasi teks khususnya berita olahraga dapat digunakan untuk mengambil isi yang paling penting dari sumber informasi yang kemudian menyajikannya kembali dalam bentuk yang lebih ringkas dari sebuah berita tanpa harus membaca seluruh isi dari berita tersebut. Namun proses kegiatan ini masih dilakukan secara manual oleh masyarakat sehingga membutuhkan waktu yang lebih lama. Karena itu diperlukan klasifikasi teks secara otomatis. Klasifikasi teks otomatis adalah mengambil isi yang paling penting dari sumber yang kemudian menyajikannya kembali dalam bentuk yang lebih ringkas dengan menjaga informasinya. Melihat hal tersebut maka dirancanglah sistem *Machine-Learning* yang menerapkan klasifikasi teks berita otomatis dengan metode *Decision Tree* [2].

Dalam pengelompokan berita penting untuk dilakukan penelitian yaitu untuk membuktikan bahwa algoritma *Decision Tree* dapat digunakan dalam melakukan pengelompokan berita dalam olahraga berdasarkan kategori olahraga: Formula 1, MMA, Basketball, Badminton, Football, dll. Dengan tujuan utama dari penelitian ini adalah untuk mendapatkan nilai akurasi yang dihasilkan agar dapat mengorganisir informasi tersebut dengan mudah dengan melakukan klasifikasi berita secara otomatis, sehingga dokumen-dokumen tersebut dapat memiliki kategori secara otomatis sesuai dengan kategori tersebut [3].

1.2 Rumusan Masalah

Berdasarkan penjelasan latar belakang yang telah disampaikan, dapat dirumuskan masalah sebagai berikut:

1.2.1 Permasalahan

Dalam era digital saat ini, berita olahraga tersebar luas di berbagai platform online. Banyaknya informasi yang tersedia membuat pembaca kesulitan dalam menemukan berita sesuai minat dengan cepat. Selain itu, berita-berita tersebut belum dikategorikan secara otomatis ke dalam jenis cabang olahraga tertentu, sehingga pencarian informasi menjadi tidak efisien.

1.2.2 Solusi Permasalahan

Solusi yang ditawarkan dalam penelitian ini adalah membangun sistem klasifikasi otomatis menggunakan algoritma Decision Tree. Sistem ini akan menjalankan tahapan sebagai berikut:

1. Preprocessing teks: Pembersihan dan normalisasi data teks (case folding, tokenizing, stopword removal, stemming).
2. Ekstraksi fitur: Menggunakan metode TF-IDF untuk mengubah teks menjadi representasi numerik.
3. Pembangunan model klasifikasi: Menggunakan algoritma Decision Tree untuk membentuk model klasifikasi berdasarkan fitur teks.
4. Evaluasi performa model: Mengukur akurasi model klasifikasi berdasarkan data uji dan matriks evaluasi seperti confusion matrix.

Dengan solusi ini, sistem diharapkan mampu mengelompokkan teks berita ke dalam kategori olahraga seperti Formula 1, MMA, Basketball, Badminton, Football, Dll secara akurat, efisien, dan dapat diinterpretasikan dengan mudah oleh pengguna.

1.2.3 Pertanyaan Penelitian

Berapa akurasi metode *Decision Tree* dalam mengklasifikasikan teks berita olahraga?

1.3 Tujuan dan Manfaat

1.3.1 Tujuan

1. Mengevaluasi performa model dalam mengelompokkan berita ke dalam kategori yang sesuai.
2. Mengetahui efektivitas Decision Tree dalam domain teks berita olahraga.

1.3.2 Manfaat

Penelitian ini diharapkan dapat membantu masyarakat dalam menentukan berita yang sesuai dengan kemauan masyarakat serta memberikan referensi penggunaan algoritma *Decision Tree* untuk pada kasus serupa.

1.4 Batasan Masalah

1. Data yang digunakan terbatas.
2. Hanya menggunakan satu metode klasifikasi yaitu *Decision Tree*.
3. Proses klasifikasi terbatas pada kategori yang telah ditentukan di dalam dataset.
4. Tidak membahas performa metode lain (*SVM*, *Naive Bayes*, *Random Forest*, dll).

1.5 Sistematika Laporan Skripsi

Sistematika penulisan pada laporan skripsi ini adalah sebagai berikut:

BAB I : PENDAHULUAN

Berisi tentang pendahuluan dari penelitian ini. Bab ini akan membahas latar belakang permasalahan yang muncul, solusi yang diusulkan untuk mengatasi permasalahan tersebut, tujuan dan manfaat dari penelitian, Batasan-batasan yang ada dalam penelitian, serta sistematika penulisan proposal.

BAB II : KAJIAN PUSTAKA

Berisi tentang kajian pustaka yang menjelaskan mengenai materi dan metode yang dijadikan landasan teori dalam penelitian ini. Di antaranya materi yang dibahas adalah Berita Olahraga, Algoritma *Decision Tree*, Klasifikasi, Preprocessing dan penelitian terkait.

BAB III : METODE PENELITIAN

Bagian ini menjelaskan mengenai Metode Penelitian yang berisi penjelasan tentang dataset yang akan digunakan, arsitektur sistem, tahapan pengujian, skenario uji coba.

BAB IV : HASIL DAN PEMBAHASAN

Membahas tentang hasil dan pembahasan, hasil dari algoritma *Decision Tree* untuk memecahkan permasalahan yang kemudian dilakukan pengujian untuk mengukur hasilnya.

BAB V : PENUTUP

Berisi kesimpulan dari implementasi yang sudah dilakukan, dan juga saran yang bertujuan agar dapat mengembangkan penelitian berikutnya.