

# BAB 1

## PENDAHULUAN

### 1.1 Latar Belakang Masalah

Dalam era digital yang semakin maju, kebutuhan akan pengolahan teks otomatis menjadi sangat penting di berbagai bidang, khususnya dalam media dan jurnalistik [2]. Salah satu tantangan yang menonjol adalah kesalahan kapitalisasi huruf dalam teks jurnalistik [2]. Kesalahan ini dapat berdampak negatif pada profesionalisme tulisan, mengurangi kredibilitas berita, dan bahkan berpotensi menyebabkan salah tafsir dalam penyampaian informasi [3]. Pedoman Ejaan yang Disempurnakan (EYD) di Indonesia memberikan aturan yang jelas mengenai penggunaan huruf kapital, namun penerapan aturan ini secara konsisten masih menjadi tantangan besar, terutama pada teks yang diproduksi dalam jumlah besar dalam waktu yang singkat [4].

Saat ini, proses proofreading manual yang digunakan untuk memastikan kualitas teks cenderung membutuhkan waktu yang lama, rentan terhadap kesalahan manusia, dan tidak efisien untuk skala industri [5]. Di sisi lain, perkembangan teknologi Pemrosesan Bahasa Alami (Natural Language Processing/NLP) menawarkan solusi yang lebih otomatis, efisien, dan dapat diandalkan untuk mendeteksi serta memperbaiki kesalahan penulisan, termasuk kesalahan kapitalisasi huruf [6]. Meskipun perangkat lunak tata bahasa dalam bahasa Inggris telah berkembang pesat, aplikasi serupa untuk bahasa Indonesia masih sangat terbatas [7]. Hal ini menunjukkan adanya kesenjangan (knowledge gap) antara kebutuhan praktis dan teknologi yang tersedia saat ini.

Untuk menjembatani kesenjangan ini, modul U-TAPIS diusulkan sebagai solusi yang dirancang khusus untuk menangani deteksi dan koreksi kesalahan kapitalisasi huruf sesuai dengan Ejaan yang Disempurnakan (EYD) [8, 9]. Modul ini memanfaatkan algoritma Random Forest sebagai inti sistemnya. Random Forest, yang merupakan metode “ensemble learning”, mampu menggabungkan berbagai pohon keputusan untuk meningkatkan akurasi prediksi. Algoritma ini terkenal karena kemampuannya dalam menangani data berdimensi tinggi, robust terhadap overfitting, dan menyediakan estimasi pentingnya fitur, sehingga sangat cocok untuk tugas klasifikasi teks seperti deteksi kesalahan kapitalisasi huruf [10, 11].

Penelitian sebelumnya mengenai U-TAPIS mengungkapkan bahwa aplikasi ini

awalnya dikembangkan sebagai alat untuk meningkatkan kemampuan bahasa para jurnalis. Dengan memanfaatkan metode pembelajaran mesin Random Forest, U-TAPIS dirancang untuk menyaring kesalahan tata bahasa Indonesia secara efisien dan akurat [12, 13]. Melalui pengembangannya, aplikasi ini tidak hanya diharapkan mampu mendeteksi kesalahan dengan lebih baik, tetapi juga berkontribusi secara signifikan dalam meningkatkan kompetensi bahasa mahasiswa jurnalisme secara keseluruhan [12, 14, 15, 16].

Target dari penelitian ini adalah mengembangkan modul U-TAPIS lebih lanjut untuk memenuhi kebutuhan dunia jurnalistik yang semakin kompleks. Modul ini diharapkan mampu mendeteksi dan memperbaiki kesalahan kapitalisasi huruf secara konsisten dan akurat, sehingga dapat mendukung praktik jurnalistik yang lebih profesional [4]. Selain itu, penelitian ini juga bertujuan untuk mengatasi keterbatasan perangkat lunak serupa dalam bahasa Indonesia, sekaligus memberikan kontribusi signifikan pada pengembangan teknologi bahasa Indonesia.

Dengan fokus pada pengembangan teknologi berbasis EYD, penelitian ini memiliki potensi besar untuk menjawab kebutuhan akan solusi otomatis dalam pengolahan teks jurnalistik. Hal ini tidak hanya mendukung kualitas tulisan, tetapi juga memperkuat kredibilitas informasi yang disampaikan kepada masyarakat luas. Oleh karena itu, topik ini penting dan layak untuk diteliti guna mendukung perkembangan teknologi bahasa Indonesia yang lebih maju dan relevan [17].

## 1.2 Rumusan Masalah

Rumusan masalah yang diangkat dalam penelitian ini meliputi poin-poin berikut:

1. Bagaimana algoritma Random Forest dapat diimplementasikan untuk mendeteksi kesalahan huruf kapital sesuai EYD5?
2. Bagaimana tingkat akurasi, *recall*, *precision*, dan *F1-score* yang dapat dicapai dalam implementasi algoritma *Random Forest* untuk mengembangkan modul U-Tapis dalam deteksi huruf kapital?

## 1.3 Tujuan Penelitian

Tujuan dari penelitian ini adalah sebagai berikut:

1. Mengembangkan model deteksi huruf kapital sesuai EYD5 menggunakan algoritma Random Forest.
2. Mengevaluasi performa model yang dikembangkan menggunakan metrik akurasi, *recall*, *precision*, dan *F1-score*.

#### **1.4 Urgensi Penelitian**

Penelitian ini memiliki urgensi yang tinggi dalam konteks pengembangan kemampuan bahasa, terutama dalam penulisan jurnalistik di Indonesia. Kualitas penulisan berita sangat penting, dan kesalahan penggunaan huruf kapital sesuai Ejaan Yang Disempurnakan (EYD) dapat merugikan kredibilitas karya jurnalistik. Dengan mengembangkan model deteksi yang akurat menggunakan algoritma Random Forest, penelitian ini berkontribusi pada peningkatan kualitas penulisan di media massa dan membantu jurnalis dalam memperbaiki kesalahan di bawah tekanan waktu. Selain itu, penelitian ini mengintegrasikan teknologi pembelajaran mesin ke dalam pendidikan bahasa bagi mahasiswa jurnalis, yang diharapkan dapat meningkatkan kompetensi mereka untuk menghadapi tantangan di dunia kerja. Hasil penelitian ini juga diharapkan dapat dipublikasikan dalam jurnal akreditasi SINTA, memberikan dampak signifikan bagi jurnalis, mahasiswa, dan masyarakat dalam meningkatkan kemampuan berbahasa dan kualitas informasi yang disajikan.

#### **1.5 Luaran Penelitian**

Penelitian ini memiliki urgensi tinggi dalam pengembangan kemampuan bahasa, khususnya dalam penulisan jurnalistik di Indonesia. Kualitas penulisan berita sangat penting, dan kesalahan penggunaan huruf kapital sesuai Ejaan Yang Disempurnakan (EYD) dapat mengurangi kredibilitas karya jurnalistik. Dengan mengembangkan model deteksi akurat menggunakan algoritma Random Forest, penelitian ini berkontribusi pada peningkatan kualitas penulisan di media massa dan memudahkan jurnalis dalam memperbaiki kesalahan di bawah tekanan waktu. Selain itu, penelitian ini mengintegrasikan teknologi pembelajaran mesin ke dalam pendidikan bahasa bagi mahasiswa jurnalis, meningkatkan kompetensi mereka untuk menghadapi tantangan di dunia kerja. Hasil penelitian ini diharapkan dapat dipublikasikan dalam jurnal akreditasi SINTA, memberikan dampak signifikan bagi

jurnalis, mahasiswa, dan masyarakat dalam meningkatkan kemampuan berbahasa dan kualitas informasi.

### **1.6 Manfaat Penelitian**

Penelitian ini diharapkan memberikan manfaat berikut:

1. Menyediakan solusi praktis untuk mendeteksi kesalahan huruf kapital di teks Bahasa Indonesia.
2. Meningkatkan pemahaman tentang penerapan aturan EYD melalui teknologi berbasis machine learning.

