

BAB 1

PENDAHULUAN

1.1 Latar Belakang Masalah

Kesalahan berbahasa merupakan masalah yang kerap dijumpai dalam berbagai konteks, baik dalam komunikasi informal sehari-hari maupun dalam penulisan formal seperti karya tulis ilmiah, dokumen hukum, maupun produk jurnalistik[1]. Kesalahan berbahasa ini dapat beragam bentuknya, mulai dari kesalahan ejaan, struktur kalimat, hingga penggunaan tanda baca yang tidak sesuai dengan kaidah yang berlaku. Dalam dunia jurnalistik, kesalahan berbahasa menjadi lebih serius karena media massa memiliki peran penting dalam menyampaikan informasi kepada publik. Kesalahan berbahasa terjadi dalam produk jurnalistik, seperti berita atau artikel dapat berdampak pada ambiguitas pembaca dan menimbulkan kesalahpahaman terhadap informasi pada berita yang diberikan. Oleh karena itu, pengecekan tata bahasa menjadi sebuah keharusan yang tidak bisa diabaikan, terutama dalam produksi teks jurnalistik yang kerap dibatasi oleh tenggat waktu.

Pengecekan tata bahasa memerlukan waktu yang cukup banyak dikarenakan perlu mengecek seluruh tata bahasa sesuai dengan Ejaan yang Disempurnakan (EYD) V[2]. Dalam situasi ini, keberadaan teknologi berbasis *Artificial Intelligence*(AI) dan *Machine Learning* (ML) menawarkan solusi yang praktis dan efisien untuk membantu pengecekan tata bahasa Indonesia secara otomatis.

Penggunaan AI dan ML dalam pengecekan tata bahasa secara otomatis memungkinkan proses deteksi kesalahan berbahasa dilakukan dengan cepat dan konsisten. Dengan algoritma yang canggih, sistem-sistem ini dapat mengidentifikasi berbagai kesalahan, mulai dari kesalahan ejaan, penggunaan tanda baca, hingga struktur kalimat yang tidak sesuai dengan kaidah yang berlaku. Selain itu, teknologi ini juga bisa terus disempurnakan seiring dengan perkembangan bahasa dan tata cara penulisan.

Untuk menerapkan teknologi AI dan ML, dibentuklah U-Tapis, sebuah aplikasi yang dibangun oleh Universitas Multimedia Nusantara untuk mengecek dan memperbaiki tata bahasa Indonesia sesuai dengan EYD V[3]. Aplikasi ini dibangun dengan harapan dapat membantu jurnalis dan editor untuk mempercepat produksi pada berita sehingga informasi dapat tersampaikan secara jelas. Per 2024,

U-Tapis telah mengembangkan modul untuk beberapa tata bahasa Indonesia, seperti U-Tapis Kata Baku, U-Tapis Kata Depan[4], U-Tapis Kata Majemuk[5], U-Tapis Kata Terikat[6], U-Tapis Saltik, U-Tapis Konjungsi, U-Tapis Peluluhan Kata, U-Tapis Kata Keterangan, U-Tapis Kalimat, Web U-Tapis I, dan Web U-Tapis II.

Pada penelitian ini, pengembangan aplikasi U-Tapis akan terus dilanjutkan dengan penelitian untuk menganalisis kesalahan pada penggunaan tanda baca. Hal ini disebabkan oleh munculnya kesalahpahaman para pembaca ketika penggunaan tanda baca disalahgunakan. Penelitian ini akan memanfaatkan *Natural Processing Language* (NLP) dan model *Recurrent Neural Network* (RNN). Keduanya dipilih karena kemampuannya dalam memproses data berurutan, yang sangat penting dalam memahami konteks penggunaan tanda baca di dalam kalimat. Penelitian ini menggunakan dataset yang terdiri dari 1.035 berita yang diambil dari situs Tribunnews.com, yang kemudian dipecah menjadi 15.783 data mentah berbentuk kalimat untuk dianalisis. Dataset ini digunakan untuk melatih model agar dapat mengenali pola kesalahan penggunaan tanda baca dalam kalimat tertentu.

Hasil dari penelitian ini diharapkan dapat memberikan kontribusi signifikan dalam pendeteksian kesalahan penggunaan tanda baca sesuai dengan tata Bahasa Indonesia. Penerapan AI dan ML melalui aplikasi seperti U-Tapis tidak hanya mampu mengurangi waktu dan tenaga yang dibutuhkan untuk melakukan pengecekan manual, tetapi juga dapat meningkatkan akurasi pengecekan tata bahasa secara konsisten. Melalui penelitian ini, teknologi yang dikembangkan juga diharapkan bisa membantu menjaga dan mempromosikan penggunaan bahasa Indonesia yang baik dan benar sesuai dengan kaidah yang berlaku, sekaligus mendorong inovasi di bidang kecerdasan buatan dalam pengolahan bahasa.

1.2 Rumusan Masalah

1. Bagaimana cara U-Tapis dapat mendeteksi kesalahan penggunaan tanda baca menggunakan *Recurrent Neural Network*?
2. Bagaimana mengimplementasikan RNN untuk mendeteksi tanda baca pada aplikasi U-Tapis?
3. Bagaimana performa dari pendeteksian kesalahan pada penggunaan tanda baca menggunakan *Recurrent Neural Network*?

1.3 Tujuan Penelitian

1. Mengembangkan modul U-Tapis untuk deteksi kesalahan penggunaan tanda baca menggunakan *Recurrent Neural Network* pada naskah yang diberikan.
2. Mengevaluasi dan menganalisis performa dari pendeteksian kesalahan pada penggunaan tanda baca menggunakan *Recurrent Neural Network* untuk naskah yang diberikan.

1.4 Urgensi Penelitian

Untuk melengkapi analisis bahasa pada aplikasi U-Tapis yang digunakan untuk keperluan menyunting naskah jurnalistik, dibutuhkan untuk pengembangan terhadap modul tanda baca. Hal ini disebabkan oleh penggunaan tanda baca yang dianalisis pada modul-modul sebelumnya yang sudah dikembangkan masih dianggap benar semua. Diharapkan dengan adanya penelitian ini, sistem dapat mendeteksi secara rinci kesalahan maupun penggunaan tanda baca yang benar.

1.5 Luaran Penelitian

1. Pendeteksian kesalahan penggunaan tanda baca pada aplikasi U-Tapis yang dikembangkan oleh Universitas Multimedia Nusantara.
2. Jurnal IEEE yang akan dipublikasikan pada lembaga jurnal nasional terakreditasi.
3. Penelitian berikutnya yang menggunakan *Recurrent Neural Network* maupun model lain yang menganalisis tingkat akurasi penggunaan model terhadap pendeteksian penggunaan tanda baca.

1.6 Manfaat Penelitian

1. Mendorong penggunaan *Natural Language Processing* untuk melakukan penyuntingan pada naskah.
2. Aplikasi U-Tapis dapat mendeteksi kesalahan penggunaan tanda baca menggunakan *Recurrent Neural Network*.
3. Membantu jurnalis untuk menyunting penggunaan tanda baca terhadap naskah yang dibentuk.