

BAB 1 PENDAHULUAN

1.1 Latar Belakang Masalah

Bahasa adalah kemampuan manusia untuk berkomunikasi dengan orang lain dengan menggunakan tanda-tanda seperti ucapan dan gerak serta mempunyai fungsi penting dalam kehidupan manusia, sebagai alat komunikasi, tanda identifikasi, dan tanda stratifikasi, sosial, dan lain-lain. Bahasa merupakan alat komunikasi yang bertujuan untuk memberikan pesan dalam bentuk ekspresi dalam berbagai aktivitas pada situasi atau kondisi tertentu[1]. Dalam hal ini, ekspresi mengacu pada unsur segmental dan suprasegmental, baik secara linguistik maupun motorik sehingga sebuah kalimat bisa berfungsi sebagai alat atau sarana komunikasi untuk menyampaikan pesan yang berbeda apabila disampaikan dalam ekspresi yang berbeda[2]. Implementasi dari kemahiran berbahasa ini adalah kemahiran dalam beretorika, baik beretorika dalam berbicara ataupun menulis. Dalam hal ini, retorika sebagai kemahiran dalam mengolah bahasa secara efisien, yaitu berupa kemampuan dalam membawa emosional pendengar atau pembaca (pathos), kemampuan dalam berkarakter atau niat baik (ethos), dan bukti logis (logos) sehingga pesan yang disampaikan melalui media tulis atau lisan dapat memengaruhi pendengar atau pembaca[3]. Bahasa merupakan aspek penting ketika kita ingin mengembangkan empat keterampilan bahasa, yaitu berbicara, membaca, menyimak, dan menulis[4]. Dengan menguasai keterampilan-keterampilan tersebut, kemampuan berkomunikasi dengan orang lain akan lebih baik.

Indonesia mempunyai bahasanya sendiri yaitu bahasa Indonesia. Bahasa Indonesia telah tercantum dalam UUD pasal 36 sebagai bahasa negara yang dinyatakan dengan "Bahasa Negara adalah Bahasa Indonesia"[5]. Sudah menjadi suatu keharusan bagi rakyat Indonesia untuk menggunakan bahasa Indonesia dengan baik dan benar. Oleh karena itu, ada aturan baku tentang penggunaan bahasa Indonesia, salah satunya adalah penggunaan kata. Terdapat banyak jenis kata dalam bahasa Indonesia, salah satunya adalah kata majemuk. Kata majemuk merupakan gabungan antara dua kata atau lebih yang membentuk sebuah makna baru. Unsur-unsur yang membentuk kata majemuk akan hilang hakikat kekataannya karena strukturnya berada dalam kesatuan gabungan itu, begitu pula hakikat kata majemuk akan hancur jika disisipkan suatu kata di tengah-tengah kata majemuk

tersebut[6].

Kesalahan penulisan kata majemuk masih sering terjadi, terlebih lagi karena kata majemuk merupakan kata gabungan dari dua kata atau lebih yang mengandung satu makna atau pengertian baru[7]. Jika dalam sebuah tulisan banyak kesalahan yang dilakukan tanpa memperhatikan ejaan yang baik dan benar, maka isi tulisan tidak dapat tersampaikan kepada pembaca dengan jelas dan tepat[8]. Seringkali penulis keliru dalam menulis kata majemuk karena penulisan kata majemuk yang perlu digabung atau dipisah. Tidak semua kata majemuk terpisah menjadi dua kata, ada juga yang menjadi satu kata yang digabung, seperti kata "olahraga", "pancaindra", dan lain-lain. Selain itu, ada juga penulis yang menggabungkan kata majemuk yang seharusnya dipisah, seperti kata "terima kasih", "orang tua", dan lain-lain. Karena masih banyaknya kesalahan dalam penulisan kata majemuk ini, dibuatlah sebuah aplikasi untuk mengoreksi kesalahan pada kata majemuk dengan menggunakan teknologi untuk bisa mendeteksi suatu kata yang dikenal dengan *Natural Language Processing(NLP)*[9].

Kata majemuk sendiri memiliki pola dalam terbentuknya kata majemuk, seperti gabungan kata benda dengan kata sifat, kata sifat dengan kata benda, kata benda dengan kata benda, kata benda dengan kata kerja, kata bilangan dengan kata benda, kata kerja dengan kata kerja, dan kata sifat dengan kata sifat[10].

Dari pola kata majemuk ini, program deteksi kata majemuk dibuat. Dalam istilah *programmer*, kasus seperti ini dinamakan dengan *rule based*. *Rule based* sendiri merupakan metode yang menerapkan aturan-aturan tertentu yang dalam hal ini untuk mendeteksi kata majemuk[11]. Dari *rule based* yang sudah dibuat, algoritma untuk mendeteksi kesalahan kata majemuk adalah algoritma *levenshtein distance*. Algoritma *levenshtein distance* ini memberi rumus untuk menghitung nilai kemiripan dari dokumen atau *string* yang dibandingkan[12]. Penggunaan algoritma *levenshtein distance* ini didapat dari *library fuzzywuzzy*

Penelitian ini merupakan pengembangan dari penelitian sebelumnya, yaitu U-Tapis yang sudah melakukan penelitian untuk pendeteksian kesalahan kata lainnya. Dikarenakan kata majemuk merupakan kata yang sering digunakan, penelitian untuk mendeteksi kesalahan kata majemuk ini dilakukan. Terlebih bahwa hasil dari penelitian ini nantinya akan digunakan oleh salah satu portal berita yaitu Tribunnews.

1.2 Maksud dan Tujuan Kerja Magang

Dalam melaksanakan penelitian terdapat maksud yang ingin dicapai, yaitu sebagai berikut.

- a. Untuk mengembangkan aplikasi web U-Tapis untuk mendeteksi kesalahan bahasa Indonesia
- b. Untuk Meningkatkan kemampuan diri baik dari segi *soft skill* maupun *hard skill* agar dapat bersaing dengan kebutuhan zaman sekarang
- c. Untuk menambah ilmu serta keahlian dalam bidang teknologi informasi serta menerapkan ilmu dan keahlian yang didapat dari perkuliahan

Adapun tujuan dari penelitian ini adalah untuk mengembangkan aplikasi web U-Tapis dalam hal mendeteksi kesalahan kata majemuk menggunakan algoritma *string matching*

1.3 Waktu dan Prosedur Pelaksanaan Kerja Magang

Waktu pelaksanaan penelitian ini dilaksanakan selama 22 minggu dengan 5 hari per minggunya dengan penjelasan sebagai berikut.

Periode	: 31 Juli 2023 - 30 Desember 2023
Hari kerja	: Senin - Jumat
Jam kerja	: 10.00 WIB - 18.00 WIB
Posisi	: Penelitian U-TAPIS

Prosedur dari kerja magang yang diselenggarakan selama program magang adalah sebagai berikut.

1. Penelitian dilakukan secara mandiri oleh anggota peneliti secara daring
2. Bimbingan penelitian dilakukan bersama dengan Marlinda Vasty Overbeek S.Kom, M.Kom, secara daring dan luring setiap bulannya.
3. Pertemuan dan penyusunan laporan untuk ketua dan seluruh tim peneliti, meliputi Niknik Mediyawati, S.Pd., M.Hum., Rudi Sutomo, S.Kom., M.Si., M.Kom, Samiaji Bintang Nusantara S.T., M.A., dan Marlinda Vasty Overbeek S.Kom, M.Kom dilakukan setiap 2 bulan