

Perangkat Lunak Identifikasi dan Koreksi Kesalahan Berbahasa Indonesia Pada Karya Ilmiah

Hadi Syaputra¹, Sunda Ariana², Margareta Andriani³, Suheriyatmono⁴
^{1,2,3}, Fakultas Ilmu Komputer, Universitas Bina Darma
e-mail : ¹hadisyaputra@binadarma.ac.id, ²sunda@binadarma.ac.id,
³m.andriani@binadarma.ac.id

Abstract. Penelitian ini berjudul Perangkat Lunak Identifikasi dan Koreksi Kesalahan Berbahasa Indonesia pada Karya Ilmiah. Tujuan jangka panjang penelitian ini menghasilkan perangkat lunak yang dapat mendeteksi dan mengoreksi kesalahan berbahasa pada saat menulis karya ilmiah berbahasa Indonesia. Perangkat lunak yang dihasilkan ini, berupa pendeteksian, opsi perbaikan, dan sekaligus koreksi kesalahan ejaan yang disempurnakan (EyD). Target khusus dari penelitian adalah sebagai berikut. Pertama, agar penulisan karya ilmiah dapat mengikuti kaidah penulisan bahasa Indonesia yang baik dan benar. Kedua, perangkat lunak ini juga sekaligus menjadi sumber belajar bahasa Indonesia pada saat menulis karya ilmiah karena perangkat lunak mampu memberikan petunjuk bagaimana memperbaiki kesalahan berbahasa dengan memberikan opsi-opsi perbaikan. Metode penelitian yang digunakan dalam penelitian ini adalah metode *waterfall* yang digunakan sebagai dasar merancang perangkat lunak yang akan dikembangkan untuk mendeteksi kesalahan-kesalahan berbahasa. Untuk pendeteksian kata imbuhan menggunakan algoritma Nazief dan Andriani, sedangkan mencari opsi kesalahan kata menggunakan algoritma *Levenshtein distance*. Perangkat lunak ini dirancang untuk dapat mengidentifikasi, mengoreksi dan sekaligus memperbaiki semua kesalahan berbahasa pada saat menulis karya ilmiah khususnya ejaan bahasa Indonesia. Capaian penelitian ini adalah terciptanya perangkat lunak yang dapat mendeteksi sekaligus mengoreksi kesalahan ejaan yang disempurnakan (EyD)

1 Pendahuluan

Bahasa Indonesia adalah sebagai bahasa resmi dalam dunia pendidikan. Sesuai dengan Pasal 36 Undang-undang Dasar 1945 menyatakan bahwa bahasa Indonesia sebagai bahasa Negara, mempunyai fungsi yaitu sebagai bahasa resmi negara, bahasa pengantar di lembaga-lembaga pendidikan, alat perhubungan pada tingkat nasional bagi kepentingan menjalankan roda pemerintahan dan pembangunan, alat mengembangkan budaya dan pemanfaatan ilmu pengetahuan, seni, serta teknologi modern. Sayangnya, tidak semua pemakai bahasa menggunakan bahasa Indonesia yang baik dan benar, sesuai dengan konteks pemakaiannya.

Masyarakat Indonesia, pada umumnya, menguasai minimal dua bahasa; bahasa ibu dan bahasa Indonesia. Bahasa ibu diperoleh pada lingkungan keluarga dan bersifat nonformal, sedangkan bahasa Indonesia diperoleh pada lingkungan formal,

misalnya di sekolah. Hal inilah yang menyebabkan penguasaan bahasa Indonesia sebagai bahasa kedua terpengaruh oleh bahasa ibu atau bahasa pertama (bahasa daerah), sehingga sering terjadi interferensi. Interferensi merupakan penerapan penggunaan dua bahasa secara bergantian sehingga menimbulkan kekacauan bahasa yang mengakibatkan kesalahan (Tarigan, 2011). Interferensi bahasa juga terjadi pada saat menulis karya ilmiah yang menuntut penggunaan ragam baku tulis dengan bahasa Indonesia yang baik dan benar.

Hal tersebut di ataslah yang menyebabkan masih banyaknya kesalahan tata tulis dalam menulis karya ilmiah berbahasa Indonesia. Padahal penulisan karya ilmiah menuntut penggunaan bahasa Indonesia yang baik dan benar. Andriani (2007) pernah melakukan penelitian mengenai kesalahan ejaan yang digunakan mahasiswa pada saat menulis karya ilmiah. Hasil penelitian menunjukkan bahwa mahasiswa masih banyak melakukan kesalahan di setiap unsur Ejaan yang Disempurnakan (EjD), yaitu pemakaian huruf, penulisan kata, pemakaian tanda baca, dan penulisan unsur serapan. Bahkan, merujuk pada hasil penelitian Ariana (2011), kurangnya penguasaan bahasa Indonesia yang baik dan benar juga terjadi pada karya ilmiah yang dihasilkan dosen. Dari hasil penelitian menunjukkan bahwa karya ilmiah dosen yang dimuat pada berkala ilmiah pun masih banyak kesalahan.

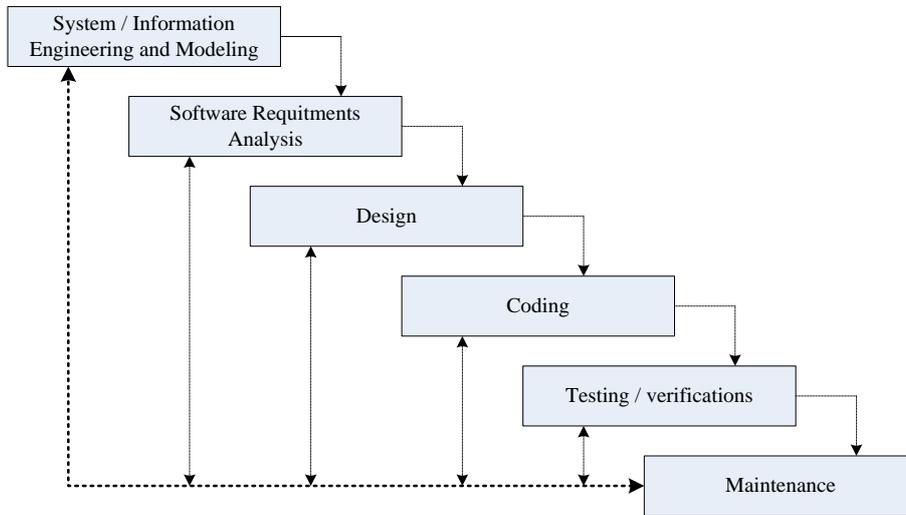
Seorang dosen harusnya menguasai, memahami, dan mengerti menggunakan tata tulis ejaan bahasa Indonesia baku ataupun kaidah-kaidah bahasa pada saat penulisan karya ilmiah mengingat mereka akan membimbing penulisan karya ilmiah mahasiswa. Di sisi lain, apabila dosen terlalu disibukkan dengan tata tulis, dikhawatirkan konten dan metodologi penelitian menjadi terabaikan. Padahal, Konten dan metodologi harusnya menjadi perhatian utama. Oleh karena itu, berdasarkan uraian di atas, maka tujuan akhir penelitian ini adalah membuat perangkat lunak yang dapat mendeteksi sekaligus mengoreksi kesalahan ejaan yang disempurnakan (EjD),

2. Landasan Teori

Penelitian ini menggunakan dua algoritma yang pertama algoritma Nazief & Adriani algoritma ini melakukan proses *stemming* yaitu mencari root dari kata. Hasil dari *stemming* akan mendapatkan kata dasar, dalam berbahasa Indonesia proses *stemming* adalah menghilangkan sufiks, prefiks, dan konfiks sehingga akan mendapatkan kata dasar yang dapat dibandingkan dengan korpus data apakah kata tersebut sesuai dengan korpus data atau tidak. Kedua algoritma *Levenshtein distance*. Algoritma ini digunakan untuk melakukan pendekatan untuk mencari kata yang sesuai apa bila terdapat kesalahan dalam penulisan kata. *Levenshtein distance* melakukan pencari berdasarkan kata awal dan kata pembanding yang ada di korpus data, yang menghasilkan opsi-opsi perbaikan kata.

3. Metode Penelitian

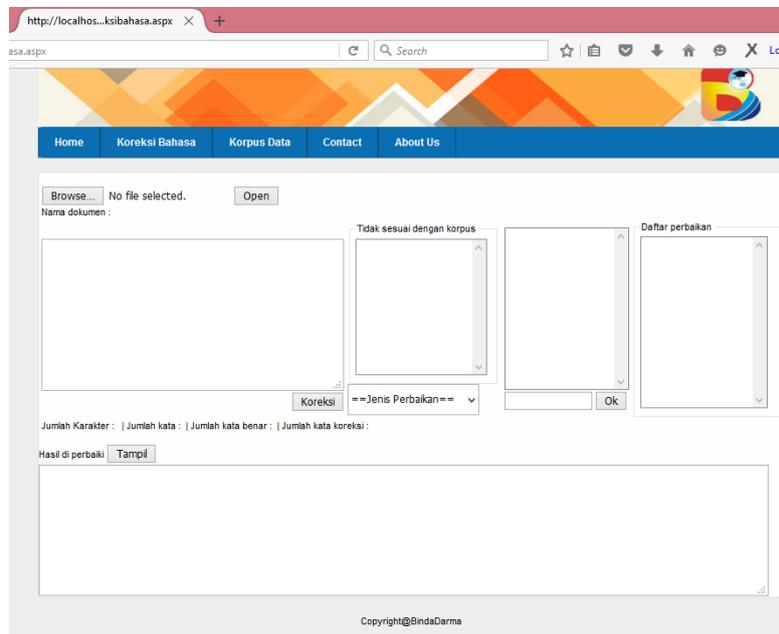
Perangkat lunak didentifikasi dan koreksi kesalahan berbahasa indonesia pada karya ilmiah ini menggunakan pendekatan model *classic life cycle* Model ini melakukan pendekatan secara sistematis dan urut mulai dari rekayasa dan pemodelan informasi/sistem, kemudian dilanjutkan dengan menganalisis kebutuhan perangkat lunak, lalu mendesain perangkat lunak, setelah desain selesai lalu mulai dengan melakukan pengkodean, setelah pengkodean akan dilanjutkan dengan ujicoba dan verifikasi, dan yang terakhir adalah pemeliharaan.



Gambar Model *Waterfall*

4. Hasil Penelitian

Adapun hasil dari rancangan perangkat lunak ini adalah sebuah Perangkat Lunak Identifikasi dan Koreksi Kesalahan Berbahasa Indonesia Pada Karya Ilmiah. Perangkat lunak ini di bangun dengan menggunakan bahasa pemrograman Visual C# ASP, dan *database* yang digunakan Microsoft SQL Server.



Perangkat lunak ini menyediakan fungsi untuk melakukan pengecekan penggunaan kata-kata yang sesuai dengan korpus data. Sistem kerja perangkat lunak ini dengan melakukan scanning kata yang ada didalam sebuah dokumen kemudian kata tersebut dicari kata dasarnya dengan menggunakan algoritma algoritma Nazief dan Andriani dan selanjutnya akan dilakukan perbandingan kata dengan kata yang terdapat pada korpus data berisikan kata-kata yang sesuai dengan EyD. Perangkat lunak ini mampu memberikan kemampuan untuk memberikan opsi perbaikan pada kata-kata yang tidak sesuai dengan EyD dengan menerapkan algoritma *Levenshtein distance*. Aplikasi koreksi yang dihasilkan dalam penelitian ini selain melakukan pengecekan kata yang sesuai dengan EyD juga dapat melakukan pengecekan penggunaan tanda baca yang tidak sesuai.

5. KESIMPULAN

Dari proses implementasi dan pengujian perangkat lunak dapat diambil beberapa kesimpulan sebagai berikut:

1. Perangkat lunak yang dibuat dapat mendeteksi kesalahan-kesalahan yang terjadi pada dokumen-dokumen karya ilmiah.
2. Perangkat lunak ini dapat menampilkan opsi-opsi perbaikan terhadap kata yang tidak sesuai dengan korpus data dan EyD.

DAFTAR PUSTAKA

- Adi Nugroho. 2005. Analisis dan Perancangan Sistem Informasi Dengan Metodologi Berorientasi Objek. Informatika. Bandung.
- Adi widya, B. M. D. 2009. Algoritma levenshtein dalam pendekatan approximate string matching. Institut Teknologi Bandung
- Andriani, Margareta. 2007. Analisis Kesalahan Ejaan karya Ilmiah: Studi Kasus Mahasiswa Nonbahasa 2007/2008 Universitas Bina Darma Palembang. *Dalam Jurnal Bina Edukasi*, vol.1 No.1 juni 2008
- Ariana, Sunda. 2011. Kesalahan Penggunaan Ejaan yang Disempurnakan dalam Karya Ilmiah Dosen Universitas Bina Darma. *Dalam Jurnal Bina Edukasi*, vol.5 No.2 Desember 2012
- Ariana, Sunda, dkk.. 2012. Prototipe Perangkat Lunak Kesalahan Berbahasa untuk Meningkatkan Kualitas Penulisan Karya Ilmiah. *Dalam Prosiding Seminar Nasional Teknologi Informasi dan Komunikasi Terapan*, ISBN: 979-26-0276-3. Semarang 2014
- Depdiknas. 2008. *Kamus Besar Bahasa Indonesia Pusat Bahasa*. Jakarta: Gramedia Pustaka Utama
- Depdiknas. 2009. *Ejaan Bahasa Indonesia yang Disempurnakan*. Jakarta: Gramedia Pustaka Utama
- Dittmar, N.. 1976. *Sociolinguistics, A Crytical Survey of Theory and Application*. London: Edward Arnold.
- Ledy Agusta. 2009. Perbandingan algoritma porter dengan algoritma nazief adriani untuk stemming dokumen teks bahasa Indonesia
- Pressman, Roger, S, 1997, *Rekayasa Perangkat Lunak: Pendekatan Praktisi*. (Edisi Satu) Yogyakarta: Andi.
- Sitompul, 2007, Analisis dan Perancangan Sistem Dengan Metodologi Berorientasi Objek: Mutiara Sumber Widya
- Sugono, Dendi. 2009. *Mahir Berbahasa Indonesia dengan Benar*. Jakarta: Gramedia Pustaka Utama.
- Tarigan, Henry Guntur. 2011. *Pengajaran Pemerolehan Bahasa*. Bandung: Angkasa.