

# PEMBELAJARAN BAHASA INDONESIA BERBASIS *SUFFIXTREE CLUSTERING* MENGGUNAKAN METODE *WEB ENGINEERING*

L. Marlinda<sup>1</sup> dan A. S. R. Lapengo<sup>2</sup>

<sup>1</sup>Teknik Komputer, AMIK Bina Sarana Informatika, <sup>2</sup>Sistem Informasi, STMIK Nusa Mandiri  
Kramat Raya 18 Jakarta Pusat

E-mail: linda.ldm@bsi.ac.id<sup>1</sup>, arif.rlg@bsi.ac.id<sup>2</sup>

## ABSTRAK

Proses belajar bahasa Indonesia saat ini masih dianggap sulit oleh siswa. Ada 4 kemampuan yang harus dikuasai yaitu mendengarkan dan membaca termasuk kategori kemampuan reseptif aktif. Sedangkan berbicara dan menulis disebut kemampuan produktif aktif. Untuk membaca dan menulis biasanya menggunakan media visual. Karena persyaratan dari penguasaan tersebut meliputi kaidah gramatikal, fonologi, morfologi, konteks kalimat dan sintaksis yang baik dan benar berdasarkan aturan yang telah ditetapkan. Sehingga dibutuhkan media dan penggunaan strategi pembelajaran yang dengan memberikan bantuan kepada siswa selama tahap awal pembelajaran dan kemudian mengurangi bantuan tersebut dan memberikan kesempatan kepada siswa untuk mengambil alih tanggung jawab yang semakin besar setelah mampu mengerjakan sendiri. Dengan dibantu media pembelajaran yang berbasis *SuffixTree Clustering* yaitu pencarian kumpulan dokumen yang memberikan hasil yang disusun berdasarkan peringkat kecocokannya. Pada penelitian ini dikembangkan suatu aplikasi pengelompokan dokumen berbasis web dengan Metode *SuffixTree Clustering*. Konsep dasarnya adalah mengelompokkan dokumen hasil pencarian ke dalam bentuk clusters berdasarkan kata. hasil penelitian ini diharapkan dapat memberikan manfaat yang sangat besar terhadap peningkatan kemampuan untuk mempermudah pengecekan penulisan atas jawaban yang dilakukan oleh siswa secara cepat dan mudah, sehingga mendapatkan hasil yang maksimal dalam pembuatan kalimat atau paragraf dalam menulis bahasa Indonesia berbasis web dan Menguji aplikasi system pengecekan kosakata Bahasa Indonesia berbasis STC yang akan dihasilkan melalui responden secara random.

**Kata Kunci:** Media Pembelajaran Bahasa Indonesia, *Suffix Tree Clustering*, *Web Engineering*

## I. PENDAHULUAN

Media pembelajaran merupakan salah satu komponen pembelajaran yang mempunyai peranan penting dalam Kegiatan Belajar Mengajar. Pemanfaatan media seharusnya merupakan bagian yang harus mendapat perhatian guru dalam setiap kegiatan pembelajaran. Oleh karena itu guru sebagai fasilitator perlu mempelajari bagaimana menetapkan media pembelajaran agar dapat mengefektifkan pencapaian tujuan pembelajaran dalam proses belajar mengajar. [1] Pada kenyataannya media pembelajaran masih sering terabaikan dengan berbagai alasan, antara lain: terbatasnya waktu untuk membuat persiapan mengajar, sulit mencari media yang tepat, tidak tersedianya biaya, dan lain-lain. Hal ini sebenarnya tidak perlu terjadi jika setiap guru mempunyai pengetahuan dan keterampilan. Dalam penelitian ini media pembelajaran yang akan digunakan menggunakan metode STC untuk mendeteksi kesalahan penulisan kosa kata dalam bahasa Indonesia.

Namun, proses belajar bahasa Indonesia masih dianggap sulit oleh siswa. penalaran untuk menjawab soal – soal esay terhadap soal karangan yang membuat siswa kurang menarik untuk mempelajarinya. kemampuan yang harus dikuasai dalam mempelajari bahasa Indonesia adalah mendengarkan, membaca dan menulis sangat membosankan bagi siswa. sehingga diperlukan suatu media visual agar pembelajaran Bahasa Indonesia menjadi menarik dan mudah.

Beberapa hal yang menyebabkan kemampuan mengarang atau menulis, menalar ataupun membaca tersebut menjadi sulit diantaranya siswa, karena belum menerapkan dan menguasai aturan kosa kata dalam hal penguasaan struktur tata bahasa dalam menulis sehingga belum menunjukkan hasil yang maksimal dan membuat motivasi siswa untuk belajar bahasa Indonesia menjadi sangat rendah dikarenakan metode pengajaran yang kurang tepat dan sebagainya.

## II. METODOLOGI PENELITIAN

### 1. Pengklasteran

Pengklasteran adalah proses pengelompokan dokumen yang sama, digabung menjadi obyek yang memiliki kedekatan yang sama. Klusterisasi merupakan teknik yang memiliki pendekatan berdasarkan tingkat persamaan di antara anggotanya [2].

## 2. Suffix Tree Clustering

Inti dari suatu hasil pencarian yang menerapkan clustering adalah penggunaan algoritma clustering. Algoritma *Suffix Tree Clustering* (STC) memiliki dua kunci utama, yaitu:

- a) Menggunakan frase sebagai dasar pembentukan cluster-nya.
- b) Menggunakan suatu definisi cluster sederhana [3,2]

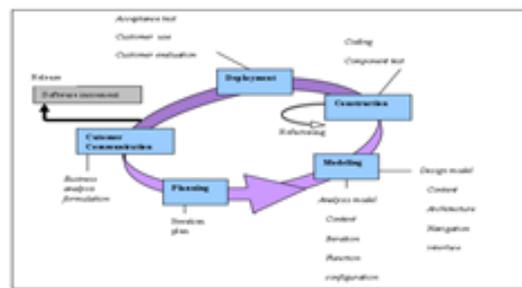
Suffix Tree Clustering (STC), memiliki tiga langkah utama [4,2]

- a. *Cleaning Dokumen* adalah dokumen yang ada, dibersihkan dan dipersiapkan untuk tahap berikutnya. Dengan dimulai dari proses analisa leksikal teks, proses penghapusan kata stopword, dan proses stemming. Pembersihan itu disebut text mining
- b. Identifikasi *Base Cluster* adalah menentukan phase tree dengan menggunakan struktur data disetiap dokumen dan dirangkai menjadi kalimat. Dengan ditentukan index dari phase sehingga membentuk score. Perhitungan jumlah dokumen yang masuk dengan jumlah kata ditunjukkan dengan persamaan  $S(B) = |B|.f(|P|)$   
 $|B|$  = jumlah dokumen di dalam base cluster B dan  $|P|$  = jumlah kata yang menyusun frase P
- c. Combining Base Cluster merupakan penanganan overlapping cluster. Dengan menghitung nilai similarity antar base cluster yang didasarkan pada jumlah dokumen yang *overlap*.

## 3. Metode Web Engineering

Metode yang digunakan dalam penelitian atau pengembangan *web* ini adalah model rekayasa *web* (*web engineering*), rekayasa *web* yang mengadaptasi rekayasa perangkat lunak dalam hal konsep dasar yang menekankan pada aktifitas teknis dan manajemen. Metode ini memerlukan pendekatan yang sistematis dan sekuensial yang mulai pada tingkat dan kemajuan sistem pada setiap tahapan. [5]

Metode *web engineering* terdapat 5 ( lima ) tahapan untuk dapat mengembangkan suatu perangkat lunak seperti Gambar 1. [6]



Gambar 1. Tahapan-tahapan Proses *Web Engineering*

Tahapan tersebut antara lain:

### 1. Customer communication

Komunikasi pengguna *web* yang akan dibangun dalam lingkungan bisnis, integrasi antara *web* dan database perusahaan.

### 2. Planning

Perencanaan proyek pengembangan aplikasi *web* yang terdiri dari pendefinisian pekerjaan dan target waktu atas pekerjaan yang telah ditentukan.

### 3. Modeling

Menjelaskan hal-hal yang diperlukan / dibutuhkan pada aplikasi yang akan dibangun, dan dapat menjawab apa yang tersirat dari hasil-hasil analisa dan pengumpulan data.

### 4. Construction

Pembangunan aplikasi *web* memadukan antara perkembangan teknologi dengan tools pengembangan *web* yang telah ada.

### 5. Deployment

Aplikasi *web* diciptakan untuk dapat berguna bagi kebutuhan pekerjaan, dapat dioperasikan oleh end-user, dan kemudian dilakukan evaluasi secara berkala.

### III. HASIL DAN PEMBAHASAN

Dalam perancangan sitem Pembelajaran Bahasa Indonesia berbasis STC (*suffix tree clustering*) menggunakan metode *web engineering*, menggunakan beberapa alat bantu untuk mempermudah dalam tahapan perancangan.

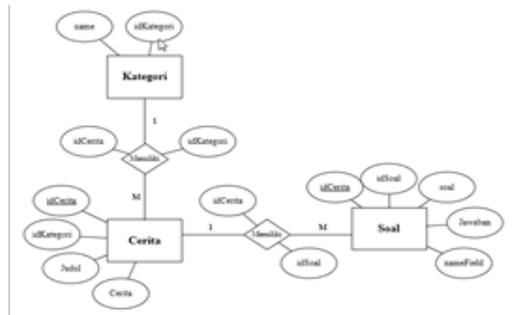
#### 1. Metodologi

Pada penelitian ini, STC yang digunakan untuk membuat cluster kata kalimat inputan dari user dengan kalimat yang ditampilkan pada halaman *web*. Dimana hasil yang keluar akan dibandingkan dengan hasil jawaban yang benar dari system. Tahapan dalam cluster hasil jawaban sebagai berikut:

- Inputan dari user berupa kalimat jawaban-jawaban di badingkan dengan kalimat-kalimat soal dan materi yang di tampilan.
- Kalimat yang sudah di filtering atau di cluster akan mengeluarkan jawaban sesuai dengan yang input dari user.

Perancangan system Pembelajaran Bahasa Indonesia yang penulis menggunakan ERD (Entity Relationship Diagram). Entity Relationship Diagram adalah suatu model untuk menjelaskan hubungan antar data dalam. Pemodelan data utama yang membantu mengorganisasikan data dalam suatu proyek kedalam entitas –entitas dan menentukan hubungan antar entitas dan basis data berdasarkan objek-objek dasar data yang mempunyai hubungan antar relasi.

Desain ERD yang ditampilkan pada Gambar 2. menunjukan hubungan entitas admin yang mengelola artikel.



Gambar 2. ERD System Pembelajaran Bahasa Indonesia

Keterangan Gambar 2 menjelaskan bahwa ERD yang digunakan menggunakan beberapa file yaitu kategori, cerita dan soal.

Selanjutnya dapat digambarkan LRS-nya seperti Gambar 3. Sedangkan tampilan sistem yang dibuat dapat dilihat pada Gambar 4, Gambar 5, Gambar 6, dan Gambar 7.



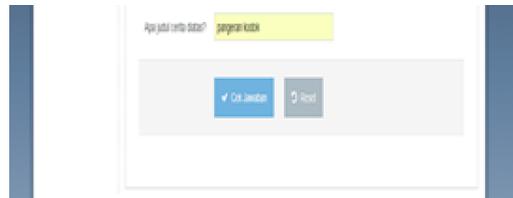
Gambar 3. LRS System Pembelajaran Bahasa Indonesia



Gambar 4. Tampilan Website System Pembelajaran Bahasa Indonesia



Gambar 5. Tampilan Soal System Pembelajaran Bahasa Indonesia



Gambar 6. Tampilan Jawaban System Pembelajaran Bahasa Indonesia

Keterangan gambar

- User mengakses *web* adalah siswa sekolah dasar dan administrator dari *web* tersebut. User siswa mengakses materi pembelajaran yang ditampilkan, serta mengakses butir latihan soal dari tiap materi pembelajaran yang di tampilkan. Administrator memiliki akses untuk menambah materi dan membuat butir soal.
- Sistem menyediakan butir soal yang diakses siswa dan melakukan penghitungan terhadap inputan jawaban dari siswa. pada proses ini dilakukan pengclusteran kalimat dengan teks dari *web*.
- Hasil pengclusteran dari kalimat dan jawaban dari siswa akan di olah sistem menggunakan STCUser siswa akan mendapatkan hasil penilaian dari butir-butir soal yang dikerjakan oleh siswa.



Gambar 7. Tampilan Hasil Akhir System Pembelajaran Bahasa Indonesia

Tampilan *web* yang penulis buat seperti gambar diatas dibangun menggunakan bahasa HTML yang disisipibahasa PHP (*Hypertext Preprocessor*). Sebagai halaman muka (*interface*) dari *web* ini adalah *index.php*. File ini merupakan induk dari template *web* dan mengatur file-file mana yang digunakan saat mengakses link-link dalam menu-menu yang ditampilkan. Di *web* ini user pengguna sebagai siswa bisa mengerjakan soal tanpa harus mendaftar. user juga bisa pengisi materi pelajaran. Siswa sebagai pengunjung *web* akan mengakses materi pelajaran dan butir-butir soal yang ditampilkan

#### IV. SIMPULAN

Pengclusteran kalimat dalam *teks mining* menggunakan berbasis STC (*suffix tree clustering*) sangat berguna untuk memberikan informasi pada dokumen yang bersifat multi redundansi dan dapat meningkatkan kecepatan pengukuran yang tepat dari kesamaan kalimat dalam hal pencarian informasi dibandingkan dengan pendekatan algoritama yang lain. Aplikasi pembelajaran yang dibuat oleh penulis diharapkan dapat membantu siswa sekolah dasar agar menyenangi pelajaran bahasa Indonesia sehingga siswa dapat lebih memahami penalaran pelajaran Bahasa Indonesia dengan mudah terutama pada soal essay.

#### V. REFERENSI

- [1] Marlinda, Harsih 2013. pembelajaran Bahasa Indonesia Berbasis *Web* Menggunakan Metode Maximum Marginal Relevance, Sesindo Bali 2013-ITS
- [2] Wicaksono, Text Mining Untuk Pencarian Dokumen Bahasa Inggris Menggunakan Suffix Tree Clustering, Jurusan Teknik Informatika, Politeknik Elektronika Negeri Surabaya, Institut

- Teknologi Sepuluh Nopember
- [3] Novan S. 2001. Implementasi Aplikasi Information Retrieval Untuk Pendeteksian dan Klasifikasi Berita Kejadian Berbahasa Indonesia
  - [4] Oren Zamir dan Oren Etzioni. 1998. *Web Document Clustering A Feasibility Demonstration*
  - [5] Turban, Mc Lean, and Wetherbe, "Information Technology for Management", Second Edition, John Wiley & Sons, Inc, 1999.
  - [6] Usman, 2013, Pengembangan system informasi online menggunakan metode web engineering, Universitas Binadarma
  - [7] Agussetia Usmaida. 2009. *Web Mining Untuk Pencarian Berdasarkan Kata Kunci Dengan Automatic Clustering*, Tugas Akhir Jurusan Teknik Informatika Politeknik Elektronika Negeri Surabaya.
  - [8] Agus Zainal Arifin, 2008. Klasifikasi Online Dokumen Berita dengan Menggunakan Algoritma Suffix *Tree* Clustering. Sesindo 2008 - ITS
  - [9] Guihong Cao, Dawei Song dan Peter Bruza. 2003. Suffix *Tree* Clustering on Post Retrieval Documents.
  - [10] Hung Chim dan Xiaotie Deng. 2007. A New Suffix *Tree* Similarity Measure for Document Clustering. Canada : IW3C2
  - [11] Oren Zamir dan Oren Etzioni. 1998. *Web Document Clustering A Feasibility Demonstration*

