

# Implementasi Text Mining Pada Penilaian Otomatis Sistem Ujian Jawaban Esai Dengan Metode Hirarki Clustering

Fris Yudha Maulana

Program Studi Teknik Informatika - S1, Fakultas Ilmu Komputer, Universitas Dian Nuswantoro Semarang

URL : <http://dinus.ac.id/>

Email : 111201005599@mhs.dinus.ac.id

**Abstract** With the development of information technology today demand a test system that is effective and efficient and able to perform the test quickly, accurately and facilitate the conduct of the examination. Currently in schools still apply test using a manual system. This manual system of course will lead to constraints in proses exam one time and cost constraints. This study will only discuss the nature essay exam for the manufacture of essay questions in the assessment of learning outcomes assessed as appropriate method, since it involves the student's ability to remember, organize, express and integrate the ideas of the students. In practice, the assessment of essays is not easy but it will use text mining method for matching string in the text, can be developed to design an automatic scoring system in the form of an essay exam answers. While clustering itself serves to categorize synonym words or words that have the same meaning. The focus of this research is the development of Text Mining in the System Implementation Exam using clustering methods. From the results are expected to facilitate the development of the students and teachers in the process of examination because it does not need to use a sheet of paper and answer many questions.

**Keywords:** *Text Mining, Automated Scoring, Hierarchical Clustering.*

## I. PENDAHULUAN

Pada akhir-akhir ini, berbagai perkembangan yang terjadi memang cukup mengesankan, khususnya dalam bidang teknologi informasi dan komunikasi. Dan perkembangan teknologi saat ini telah menyentuh hampir semua sektor. Hal ini perlu adanya peningkatan kinerja baik dari segi efektifitas dan efisiensi. Salah satu sektor yang menjadi hal penting untuk diperhatikan peningkatan kinerjanya ialah sektor pendidikan [1].

Indonesia merancang sebuah sistem pendidikan yang telah lama bertahan dan telah digunakan di sekolah-sekolah negeri maupun swasta. Seperti yang diketahui pendidikan di Indonesia dibagi ke dalam empat jenjang yaitu anak usia dini, dasar, menengah dan tinggi. Saat ini seluruh penduduk Indonesia diharuskan mengikuti program wajib belajar pendidikan dasar selama Sembilan tahun. Sedangkan siswa yang ingin melanjutkan jenjang ke berikutnya harus mengikuti proses ujian Negara. Dengan Perkembangan teknologi informasi saat ini menuntut suatu sistem pengujian yang efektif dan efisien serta mampu melakukan pengujian secara cepat, tepat dan memudahkan dalam melakukan pengujian serta penilaian itu sendiri. Pada penelitian ini akan dilakukan perancangan sistem ujian

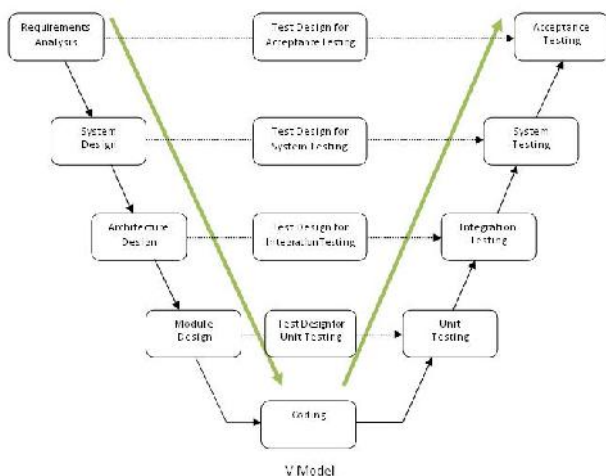
dengan memanfaatkan soal esai. Dalam pembuatan soal esai dalam penilaian hasil kegiatan belajar dinilai sebagai metode yang tepat, karena melibatkan kemampuan siswa untuk mengingat, mengorganisasikan, mengekspresikan dan mengintegrasikan gagasan yang dimiliki siswa tersebut. Dalam pelaksanaannya, melakukan penilaian esai tidaklah mudah namun dengan memanfaatkan metode untuk pencocokan *string* pada teks, dapat dikembangkan untuk merancang sistem penilaian otomatis pada jawaban ujian berbentuk esai. Pada ujian esai kebanyakan dilakukan dengan jumlah mahasiswa yang banyak, sehingga diperlukan sistem penilaian otomatis (*autograding*) pada jawaban ujian esai. Model soal dengan jawaban pendek (*short answer*) merupakan nilai penting dalam mengatur dan menyatukan ide, selain itu juga mampu menginterpretasikan data dan aplikasi dari suatu data. Namun kesulitan terbesar dari model ini adalah bagaimana subyektifitas dari masing-masing jawaban dapat dimengerti. Dalam pelaksanaannya, penelitian ini akan mengimplementasikan proses *Text mining*. *Text Mining* adalah salah satu teknik yang dapat digunakan untuk melakukan klasifikasi dimana, *text mining* merupakan variasi dari data mining yang berusaha menemukan pola yang menarik dari sekumpulan data tekstual yang berjumlah besar. *Text Mining* juga memiliki definisi yaitu menambang data yang berupa teks dimana sumber data biasanya didapatkan dari dokumen, dan tujuannya adalah mencari kata-kata yang dapat mewakili isi dari dokumen sehingga dapat dilakukan analisa

keterhubungan antar dokumen [2]. Salah satu teknik yang dikenal dalam data mining yaitu *clustering*. Pengertian *clustering* keilmuan dalam data mining adalah pengelompokan sejumlah data atau objek ke dalam *cluster* (*group*) sehingga setiap dalam *cluster* tersebut akan berisi data yang semirip mungkin dan berbeda dengan objek dalam *cluster* yang lainnya. Sampai saat ini, para ilmuwan masih terus melakukan berbagai usaha untuk melakukan perbaikan model *cluster* dan menghitung jumlah *cluster* yang optimal sehingga dapat dihasilkan *cluster* yang paling baik. Ada dua metode *clustering* yang kita kenal, yaitu *hierarchical clustering* dan *partitioning*. Metode *hierarchical clustering* sendiri terdiri dari *complete linkage clustering*, *single linkage clustering*, *average linkage clustering* dan *centroid linkage clustering*. Sedangkan metode *partitioning* sendiri terdiri dari k-means dan fuzzy k-means [3].

Berdasarkan masalah tersebut, akan dibuat suatu aplikasi penilaian otomatis pada sistem ujian jawaban esai yang berfungsi untuk member kemudahan dalam proses ujian dengan menggunakan metode hirarki clustering.

## II. METODE YANG DIUSULKAN

Model proses pengembangan atau rekayasa perangkat lunak yang digunakan dalam penelitian ini adalah V-Model. Model ini merupakan perluasan dari model waterfall. Disebut sebagai perluasan karena tahap-tahapnya mirip dengan yang terdapat dalam model waterfall. Jika dalam model waterfall proses dijalankan secara linear, maka dalam model V proses dilakukan bercabang. Dalam model V ini digambarkan hubungan antara tahap pengembangan software dengan tahap pengujiannya



Gambar 1 Aliran Proses Metodologi V-Model

Adapun penjelasan mengenai tahapan metodologi dalam pengembangan sistem ini adalah:

### 1. Requirement Analysis & Acceptance Testing

Tahap Requirement Analysis sama seperti yang terdapat dalam model waterfall. Keluaran dari

tahap ini adalah dokumentasi kebutuhan pengguna. Acceptance Testing merupakan tahap yang akan mengkaji apakah dokumentasi yang dihasilkan tersebut dapat diterima oleh para pengguna atau tidak.

### 2. System Design & System Testing

Dalam tahap ini analisis sistem mulai merancang sistem dengan mengacu pada dokumentasi kebutuhan pengguna yang sudah dibuat pada tahap sebelumnya. Keluaran dari tahap ini adalah spesifikasi software yang meliputi organisasi sistem secara umum, struktur data, dan yang lain. Selain itu tahap ini juga menghasilkan contoh tampilan window dan juga dokumentasi teknik yang lain seperti Entity Diagram dan Data Dictionary.

### 3. Architecture Design & Integration Testing

Sering juga disebut High Level Design. Dasar dari pemilihan arsitektur yang akan digunakan berdasar kepada beberapa hal seperti: pemakaian kembali tiap modul, ketergantungan tabel dalam basis data, hubungan antar interface, detail teknologi yang dipakai.

### 4. Module Design & Unit Testing

Sering juga disebut sebagai Low Level Design. Perancangan dipecah menjadi modul-modul yang lebih kecil. Setiap modul tersebut diberi penjelasan yang cukup untuk memudahkan programmer melakukan coding. Tahap ini menghasilkan spesifikasi program seperti: fungsi dan logika tiap modul, pesan kesalahan, proses input-output untuk tiap modul, dan lain-lain.

### 5. Coding

Dalam tahap ini dilakukan pemrograman terhadap setiap modul yang sudah dibentuk.

## III. IMPLEMENTASI

implementasi yang dibahas pada bab ini adalah implementasi *text mining*, yaitu proses penambangan data berupa teks dimana sumber data biasanya didapatkan dari dokumen, dan tujuannya adalah mencari kata-kata yang dapat mewakili isi dari dokumen sehingga dapat dilakukan analisa keterhubungan antar dokumen. Proses penambangan teks ini secara umum dilakukan dengan beberapa tahapan yaitu :

### 1. Tokenizing

Tahap tokenizing adalah tahap pemotongan string input berdasarkan tiap kata yang menyusunnya. Berikut proses tokenizing atau pemotongan tiap kata

Id	Token
<input type="checkbox"/>	1 hubung
<input type="checkbox"/>	1 bagai
<input type="checkbox"/>	1 jenis
<input type="checkbox"/>	1 komputer
<input type="checkbox"/>	1 bentuk
<input type="checkbox"/>	1 jaring
<input type="checkbox"/>	1 komputer
<input type="checkbox"/>	1 global
<input type="checkbox"/>	1 komunikasi
<input type="checkbox"/>	1 gunakan
<input type="checkbox"/>	1 jalur
<input type="checkbox"/>	1 telekomunikasi
<input type="checkbox"/>	2 hubung
<input type="checkbox"/>	2 bagai
<input type="checkbox"/>	2 jenis
<input type="checkbox"/>	2 komputer
<input type="checkbox"/>	2 bentuk
<input type="checkbox"/>	2 jaring

Gambar 2 Hasil dari tokenizing

2. Filtering

Tahap mengambil kata-kata penting dari hasil token. Biasanya menggunakan *stop list* (membuang kata-kata yang kurang penting) atau *word list* (menyimpan kata-kata penting). Pada sistem ini menggunakan *word list*, tampilan hasil filtering sebagai berikut :

Id	Filter
<input type="checkbox"/>	1 adalah
<input type="checkbox"/>	2 sebuah
<input type="checkbox"/>	3 di
<input type="checkbox"/>	4 ke
<input type="checkbox"/>	5 dari
<input type="checkbox"/>	6 pada
<input type="checkbox"/>	7 ini
<input type="checkbox"/>	8 dulu
<input type="checkbox"/>	9 agar
<input type="checkbox"/>	10 karena
<input type="checkbox"/>	11 yang
<input type="checkbox"/>	12 dan
<input type="checkbox"/>	13 dalam
<input type="checkbox"/>	14 dengan
<input type="checkbox"/>	15 antara
<input type="checkbox"/>	17 )
<input type="checkbox"/>	18 (
<input type="checkbox"/>	19 ,

Gambar 3 Hasil dari Filtering

3. Stemming

Tahap *stemming* adalah tahapan dimana mencari kata dasar dengan cara menghilangkan imbuhan awalan dan akhiran. Hasil dari *stemming* adalah sebagai berikut :

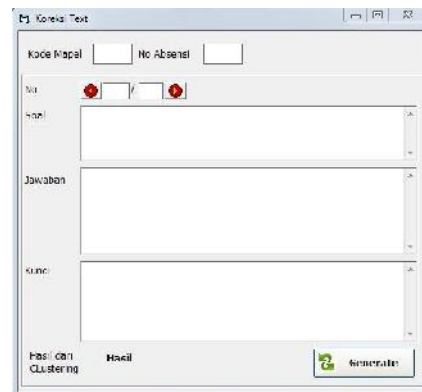
Id	Stem	JmlChar	Jenis
<input type="checkbox"/>	1 di	2	1
<input type="checkbox"/>	2 ber	3	1
<input type="checkbox"/>	3 ter	3	1
<input type="checkbox"/>	4 men	3	1
<input type="checkbox"/>	5 per	3	1
<input type="checkbox"/>	6 me	2	1
<input type="checkbox"/>	7 mer	3	1
<input type="checkbox"/>	8 kan	3	3
<input type="checkbox"/>	9 lah	3	3
<input type="checkbox"/>	10 meng	4	1
<input type="checkbox"/>	11 an	2	3
<input type="checkbox"/>	12 nya	3	3
*	(NULL) (NULL)	(NULL)	(NULL)

Gambar 4 Hasil dari stemming

• Implentasi Admin Interface



Gambar 5 Admin Interface Halaman Home

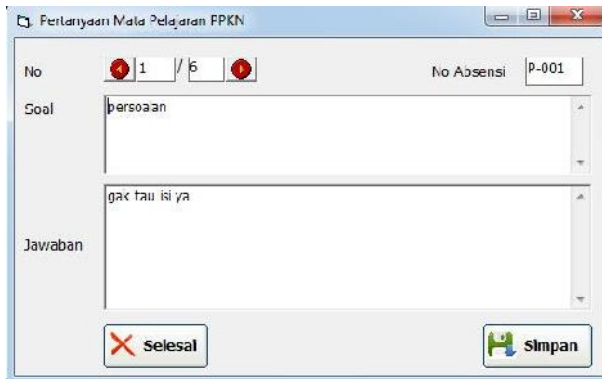


Gambar 6 Halaman Koreksi

• Implementasi User Interface



Gambar 7 Halaman absen User Interface



Gambar 8 Halaman Soal Ujian

#### IV. HASIL & PEMBAHASAN

- Black Box Testing

Pengujian sistem ini dilakukan dengan menggunakan metode pengujian *blackbox*. pengujian black box, mengevaluasi hanya dari tampilan luarnya (interface nya), fungsionalitasnya. Tanpa mengetahui apa sesungguhnya yang terjadi dalam proses detilnya (hanya mengetahui input dan output). Berikut Hasil Pengujian BlackBox :

Table 1 Pengujian Blackbox Halaman Kuis

No	Keadaan Awal	Skenario	Yang di harapkan	Hasil Uji
1	Halaman awal kuis peserta	Masukan nomor absen atau peserta yang telah terdaftar	User atau Peserta bisa mengikuti ujian yang telah di sediakan	Tercapai
2		Mengisi semua jawaban dari soal-soal yang tersedia	User dapat menekan tombol selesai tanpa adanya peringatan	Tercapai
3	Halaman Soal kuis peserta	Mengisi jawaban setiap nomer lalu pilih simpan	Muncul pemberitahuan "Data telah tersimpan"	Tercapai
4		Mengosongkan jawaban pada salah satu soal lalu pilih selesai	Muncul peringatan "masih ada soal yang kosong harap dijawab terlebih dahulu"	Tercapai

Table 2 Pengujian Blackbox Login Admin

No	Keadaan Awal	Skenario	Yang di Harapkan	Hasil Uji
1	Halaman login Admin	Admin login menggunakan username dan password yang sesuai	Admin dapat masuk ke halaman kuis admin	Tercapai
2		Admin login dengan username dan password yang tidak sesuai	Admin tidak dapat masuk ke halaman kuis admin	Tercapai
3		Admin login dengan username yang tidak sesuai	Admin tidak dapat masuk ke halaman admin dan muncul peringatan "Gagal Login"	Tercapai
4	Halaman login Admin	Admin login dengan password yang tidak sesuai	Admin tidak dapat masuk ke halaman admin dan muncul peringatan "Gagal Login"	Tercapai

Tabel 3 Pengujian Blackbox Halaman Mapel

No	Halaman Awal	Skenario	Yang di Harapkan	Hasil Uji
1	Halaman Daftar tambah Mapel	Admin mengisi seluruh form mapel	Berhasil di simpan	Tercapai
2	Halaman Daftar tambah Mapel	Admin tidak mengisi salah satu form mapel	Muncul peringatan	Tercapai
3	Halaman Daftar edit mapel	Admin mengkosongkan form mapel	Muncul peringatan	Tercapai
4	Mapel	Admin merubah salah satu form mapel	Berhasil di simpan	Tercapai
5	Halaman daftar hapus mapel	Menghapus daftar mapel	Muncul peringatan "Apakah yakin ingin menghapus" klik Ya	Tercapai

Tabel 4 Pengujian Blackbox Daftar Petugas

No	Halaman Awal	Skenario	Yang di Harapkan	Hasil Uji
1	Halaman Daftar tambah Petugas	Admin mengisi seluruh form petugas	Berhasil di simpan	Tercapai
2	Halaman Daftar tambah Petugas	Admin tidak mengisi salah satu form petugas	Muncul peringatan	Tercapai
3	Halaman Daftar edit petugas	Admin mengkosongkan form petugas	Muncul peringatan	Tercapai
4	Petugas	Admin merubah salah satu form petugas	Berhasil di simpan	Tercapai

Tabel 5 Pengujian Blackbox Daftar Peserta

No	Halaman Awal	Skenario	Yang di Harapkan	Hasil Uji
1	Halaman Daftar tambah Peserta	Admin mengisi seluruh form peserta	Berhasil di simpan	Tercapai
2	Halaman Daftar tambah Peserta	Admin tidak mengisi salah satu form peserta	Muncul peringatan	Tercapai
3	Halaman Daftar edit peserta	Admin mengkosongkan form peserta	Muncul peringatan	Tercapai
4	Peserta	Admin merubah salah satu form peserta	Berhasil di simpan	Tercapai
5	Halaman daftar hapus peserta	Menghapus daftar peserta	Muncul peringatan "Apakah yakin ingin menghapus" klik Ya	Tercapai

## V. PENUTUP

Proses pengkoreksian atau penilaian menggunakan *text mining* lebih cepat dan tidak memerlukan banyak waktu, karena jawaban akan langsung dicocokkan dengan kunci jawaban yang tersedia. Sistem Ujian ini dirancang dan dibangun sangat sederhana karena untuk memudahkan siswa atau pengguna dalam mengoperasikannya.

## REFERENCES

- [1] Adrifina, Arrummaisha., Juwita Utami Putri, dan I wayan Simri W. 2008. “*Pemilihan Artikel Berita Dengan Text Mining*”. Seminar Ilmiah Nasional Komputer dan Sistem Intelijen (KOMMIT 2008).
- [2] Fiktori, Lambang Putera (2014). Sistem Baru UN.From <http://suaraguru.wordpress.com/2014/03/22/sistem-baru-un/>, 11 juli 2014.
- [3] Muhammad Zakiya Nafik, Indriati, ST., M.Kom dan Akmad Ridok, M.Kom, 2012, “*Sistem Penilaian Otomatis Jawaban Esai Menggunakan Algoritma Levenshtein Distance*”. Universitas Brawijaya Malang.
- [4] Alfina, Tahta, Budi Santosa dan Ali Ridho Barakbah, 2012, “*Analisa Perbandingan Metode Hierarchical Clustering, K-means dan Gabungan Keduanya dalam Cluster Data*”. Institut Teknologi Sepuluh Nopember (ITS).vol. 1, No 1.
- [5] Kurniawan Bambang, Syahril Effendi dan Opim Salim Sitompul, 2012, “*Klasifikasi Konten Berita Dengan Metode Text Mining*”. Universitas Sumatera Utara. Volume 1. No. 1.
- [6] Study in the USA (2014).Memahami Sistem Pendidikan Amerika.From <http://studyusa.com/id/a/258/memahami-sistem-pendidikan-amerika>, 14 juli 2014
- [7] Kompasiana (2014).”*Virtual Exam Gaya Ujian Di Era Teknologi Modern*”.from <http://edukasi.kompasiana.com/2014/03/03/virtual-exam-gaya-ujian-di-era-teknologi-modern--635990.html>, 14 juli 2014.
- [8] Indranandita, Amalia., Budi Susanto dan Antonius Rachmat C. 2008. “*Sistem Klasifikasi Dan Pencarian Jurnal dengan Menggunakan Metode Naïve Bayes dan Vektor Space Model*”.10 Jurnal Informatika. Volume 4.No 2.
- [9] Raymond J. Mooney. CS (2006), “*Machine Learning Text Categorization*”. University of Texas : Austin
- [10] Wakhidah, Nur. 2010. “*Clustering Menggunakan K-MEANS Algorithm (K-MEANS Algorithm Clustering)*”. Universitas Semarang.
- [11] Mulyanto, Aunur R. 2008.Rekayasa Perangkat Lunak.Jakarta: Direktorat Pembinaan Sekolah Menengah Kejuruan
- [12] Dharwiyanti, S.2003. Pengantar Unified Modelling Language (UML). Ilmu Komputer.