

ANALISA DAN INPLEMENTASI

Boy Gunawan Sirait¹, Dhinta Darmantoro², Yanuar Firdaus A.w.³

¹Teknik Informatika, Fakultas Teknik Informatika, Universitas Telkom

Abstrak

Banyaknya universitas di Indonesia saat ini telah menghasilkan banyak judul TA. Pada umumnya sebelum mengajukan judul TA, maka mahasiswa terlebih dahulu memeriksa apakah judul TA tersebut sudah pernah ada. Untuk menghindari kesamaan judul TA, sistem ini dibangun berbasis web sehingga pihak yang berkepentingan dapat memeriksa judul TA yang diajukan dengan judul TA yang ada dari seluruh Universitas secara online

Metoda yang digunakan dalam implementasi sistem yang akan dibuat adalah Latent Semantic Indexing (LSI). LSI adalah algoritma yg menentukan apakah hasil pencarian relevan dengan apa yg dicari dengan cara memproses data berdasarkan jumlah kata. LSI mampu menentukan relevansi sebuah kalimat dengan membandingkan tiap kalimat yang sudah ada didalam database. Semakin banyak kata dalam dokumen yang sama dengan kata-kata dalam query, semakin tinggi peringkat kesesuaian dokumen tersebut. Tujuan dari LSI adalah mendapatkan suatu pemodelan yang efektif untuk merepresentasikan hubungan antara kata kunci dan dokumen yang dicari. Berdasarkan hasil analisis dan pengujian sistem yang dilakukan dapat diambil beberapa kesimpulan yaitu dengan bertambahnya jumlah data maka waktu yang dibutuhkan untuk pembentukan matriks term dokumen dan waktu pencarian semakin besar, Nilai persentase rank k yang terbaik dalam proses pencarian query adalah 20% dari jumlah data yang diinputkan, perubahan rank k tidak mempengaruhi waktu pencarian query dimana dengan jumlah data tertentu waktu pencarian cenderung konstan, perubahan query tidak merubah nilai rank k terbaik dan tidak berpengaruh besar terhadap waktu pencarian.

Kata Kunci : LSI, judul TA, relevan, precision, recall

Abstract

Nowdays many University in Indonesia have result many final task, Genarally before final task proposed, firt University student have to chek the final task before. This system is designed based on web to avoid final task similarity, so who intersted on can check final task propose among final task which ever had in all university by online.

The method used in implementation system is Latent Semantic Indexing (LSI). LSI is a algorithm LSI adalah algoritma yg menentukan apakah hasil pencarian relevan dengan apa yg dicari dengan cara memproses data berdasarkan jumlah kata

LSI can determine a relevant statement to compare with other statement in database. Searching more accurate if more word may have in a document. Purpose of LSI is find a efectif model to representation a relation between key word and document that is searched.

Beside on the result of analysis and testing. The conclusions are increase of data can make increase of searching time. The best persentase of rank k when searching query 20 % from total data, change of rank k not to be influenced searching time, change of query can not change the best rank k and can not to be influential concern to searcing time.

Keywords : LSI, the title of TA, relevant, precision, recall

1. PENDAHULUAN

1.1 Latar Belakang

Tugas akhir merupakan salah satu kewajiban yang harus ditempuh oleh mahasiswa agar dapat dinyatakan lulus sebagai sarjana Strata-1. Tugas akhir merupakan karya tulis ilmiah yang dapat berupa: penelitian lapangan, penelitian pustaka, pembuatan software/hardware sesuai minat dan bidang studi yang ditempuh. Tujuan utama dari tugas akhir adalah membekali mahasiswa agar dapat menuangkan kemampuan teoritisnya sesuai dengan minat yang telah ditekuninya dalam bentuk tulisan ilmiah. Tetapi yang terjadi saat ini banyak mahasiswa yang menempuh jalan pintas dengan melakukan plagiat tugas akhi, bahkan tidak sedikit tugas akhir yang diperjualbelikan. Untuk menghindari hal tersebut tugas akhir ini dibuat sehingga pihak yang berkepentingan dapat memutuskan kelayakan tugas akhir yang diajukan.

Metoda yang digunakan dalam implementasi sistem yang akan dibuat adalah *Latent Semantic Indexing* (LSI). LSI adalah algoritma yg menentukan apakah hasil pencarian relevan dengan apa yg dicari dengan cara memproses data berdasarkan jumlah kata. LSI mampu menentukan relevansi sebuah kalimat dengan membandingkan tiap kalimat yang sudah ada didalam *database*.

Dengan penggunaan metode LSI, sistem akan menganalisis kata kuncinya berdasarkan pencocokkan kata secara leksikal. Judul tugas akhir yang dicari tidak hanya kata kuncinya saja seperti pada algoritma pada umumnya, tetapi kata-kata yang berhubungan dengan kata kunci juga dicari sehingga hasil pencarian sedekat mungkin dengan judul yang diinputkan untuk diperiksa kemiripannya. Dengan sistem ini diharapkan pihak yang berkepentingan dapat lebih akurat memutuskan kelayakan tugas akhir yang diajukan.

1.2 Perumusan Masalah

Permasalahan yang dijadikan objek penelitian dan analisa dan tugas akhir ini adalah bagaimana suatu aplikasi web dapat mendeteksi kesamaan judul tugas akhir yang diajukan oleh seseorang mahasiswa dengan tugas akhir sebelumnya yang telah dibuat

Pembahasan dalam tugas akhir ini dengan batasan sebagai berikut:

1. Menggunakan salah satu teknik *Information Retrieval* yaitu *Singular Value Decomposition* (SVD)
2. Seluruh pengambilan keputusan seluruhnya diserahkan kepada admin, sistem hanya memberikan persentase kemiripan.
3. Tidak menangani sinonim dan kata berimbuhan.

4. Data yang digunakan adalah judul TA Sekolah Tinggi Teknologi Telkom Departemen Teknik Informatika di inputkan secara manual.

1.3 Tujuan

Tujuan yang ingin dicapai dari penulisan tugas akhir ini adalah:

1. Merancang sistem berbasis web untuk memberikan persentase(%) kemiripan suatu judul TA dengan judul TA yang ada dalam *database* dengan menggunakan metoda *Latent semantic indexing* (LSI).
2. Menganalisis akurasi dan waktu pencarian dari aplikasi dibangun.

1.4 Metodologi Penyelesaian Masalah

Metodologi yang digunakan untuk menyelesaikan tugas akhir ini adalah:

1. Studi literatur
Melakukan studi literatur dengan melakukan studi pustaka dan mencari referensi tentang konsep metode *Latent Semantic Indexing* dan web *programming*.
2. Analisa dan Perancangan Sistem
Menganalisis kebutuhan sistem. Dari hasil analisis tersebut dibuat perancangan aplikasinya.
3. Implementasi
Mengimplementasikan hasil dari perancangan sistem dengan membangun aplikasi sistem pendeteksi kesamaan judul TA dengan menggunakan teknik *Singular Value Decomposition* (SVD).
4. Analisa dan Pengujian
Melakukan pengujian terhadap parameter performansi sistem, hasil pengujian kemudian dianalisis untuk mendapatkan kesimpulan akhir.
5. Kesimpulan Akhir
6. Penyusunan laporan tugas akhir dan kesimpulan.

Telkom
University

5. PENUTUP

5.1 Kesimpulan

Berdasarkan hasil analisis dan pengujian sistem yang dilakukan dapat diambil beberapa kesimpulan yaitu :

1. Dari pengaruh jumlah data yang digunakan
Dengan bertambahnya jumlah data maka waktu yang dibutuhkan untuk pembentukan matriks term dokumen dan waktu pencarian semakin besar.
2. Dari pengaruh perubahan *rank k* yang digunakan
Nilai persentase *rank k* yang terbaik dalam proses pencarian *query* adalah 20% dari jumlah data yang diinputkan, perubahan *rank k* tidak mempengaruhi waktu pencarian *query* dimana dengan jumlah data tertentu waktu pencarian cenderung konstan, untuk mendapatkan nilai precision yang besar maka nilai *k* sistem diset besar, sedangkan untuk mendapatkan recall yang besar maka nilai *k* pada sistem diset kecil.
3. Dari pengaruh perubahan *query*
Perubahan *query* tidak merubah nilai *rank k* terbaik sesuai yang terdapat pada point 2, perubahan *query* tidak berpengaruh besar terhadap waktu pencarian.

5.2 Saran

Sebagai tindak lanjut dari penelitian yang dilakukan untuk meningkatkan sistem ini, ada beberapa saran yang bisa disampaikan yaitu:

1. Untuk pengembangan lebih lanjut diharapkan sistem yang dibuat dapat menangani sinonim dari suatu term.
2. Agar didapatkan sistem yang terbaik diharapkan dilakukan pengujian secara kompleks seperti: sistem database untuk menangani data yang besar, kemampuan suatu *resource* komputer.

Telkom
University

DAFTAR PUSTAKA

- [1] Castagnetto, Jesus, et al.,eds. *Professional PHP Programming*. PDF
- [2] Chakrabarti. *Mining the Web*. 2003
- [3] Garcia. E, Singular Value Decomposition (SVD) A Fast Track Tutorial, <http://www.milista.com>.
- [4] Haryanto, Steven. *PHP: kumpulan resep pemrograman*. Jakarta: Dian Rakyat, 2004
- [5] <http://indo-problogger.blogspot.com/2007/latent-semantic-indexing-2.html>
- [6] <http://www.nameslot.com/Latent-Semantic-Indexing-This-Is-New-t-1071.html#pid3696/> diakses tanggal 20 Februari 2007 pukul 22.00
- [7] <http://www.search-marketing.info/search-enginehistory/> diakses tanggal 20 Februari 2007 pukul 22.00
- [8] <http://www.stanford.edu/class/cs276/> diakses tanggal 25 Februari 2007 pukul 22.00
- [9] <http://www.cs.utk.edu/~letsche> diakses tanggal 25 Februari 2007 pukul 20.00
- [10] Salton, G and M.I McGill. 1984. Introduction to modern information retrieval. New York: McGraw-Hill.
- [11] Sullivan, Danny. Major search engines and directories. <http://www.searchenginewatch.com>.
- [12] Versi draft: C. Manning, P. Raghavan, and H. Schütze. *Introduction to Information Retrieval*. 2006
- [13] Welling, Luke, and Thomson, Laura. *PHP and Mysql Web Development*. PDF