

# PENCARIAN KEMIRIPAN JUDUL TUGAS AKHIR MAHASISWA DENGAN MENGGUNAKAN METODE SINGLE LINKAGE HIERARCHICAL

Rizal Tjut Adek<sup>1</sup>, Miftahul Jannah<sup>2</sup>  
1,2)Teknik Informatika Universitas Malikussaleh  
Jl. Batam No. 8, Kampus Bukit Indah, Lhokseumawe  
Email: rizal@unimal.ac.id

## ABSTRACT

*The search for similarity in the final project title based on the theme of the informatics engineering department using the single linkage hierarchical method is a method to determine the similarity or closeness of the abstract and the title of the final project between the input entered by the admin and the abstract and the title of the final project that has been done or is already in the department of informatics engineering Malikussaleh University with clustering techniques. In the abstract data of the final project that has been done or already done, the clustering process is carried out by using Single Linkage Hierarchical Method (SLHM) until six clusters are formed according to the fields in the informatics engineering department. Then the existing input passes the text mining process with six clusters formed. Then the matching process is carried out between test data or new data with existing data with members of the cluster. The data used to form data clustering are abstract data of informatics engineering final assignments at Malikussaleh University in 2010 - 2015, while the abstracts inputted are new abstracts to find out which assets are included in which categories are based on clustering already in the database. The results of the experiment 60 test data abstract the percentage of match success in the multimedia category by 100%, the programming category by 100%, the image processing category by 87.5%, the pattern recognition category by 11.11%, while in the network and data mining categories not found funny. And on the user page, the results in the form of the final assignment title in the database are based on the theme of the final assignment that is entered by the user.*

**Keyword** : text mining, single linkage hierarchical, clustering, abstrak dan judul tugas akhir mahasiswa.

## PENDAHULUAN

Dalam bidang kajian pencarian informasi (*Information Retrieval*) sering dimanfaatkan untuk pencarian dokumen dengan tingkat kemiripan sangat tinggi, yang juga bisa diterapkan dalam bidang akademis (Aggarwal & Rendy, 2013; Hachaj & Ogiela, 2017), misalnya bagi mahasiswa dalam pencarian kemiripan judul Tugas Akhir. Tugas akhir tentu

hal yang sudah biasa didengar oleh mahasiswa terutama mahasiswa tingkat akhir. Karena, untuk memenuhi persyaratan kelulusan program pendidikan jenjang S-1 mahasiswa biasanya diharuskan membuat sebuah laporan tugas akhir berbentuk karya tulis ilmiah.

Sejauh ini pencarian informasi khususnya dalam pencarian kemiripan

judul tugas akhir masih dilakukan secara manual dimana mahasiswa yang ingin mengetahui judul-judul terdahulu harus mengunjungi perpustakaan untuk melihat dan membaca satu persatu judul skripsi yang sudah ada sehingga masih kurang terarah dan kurang efektif.

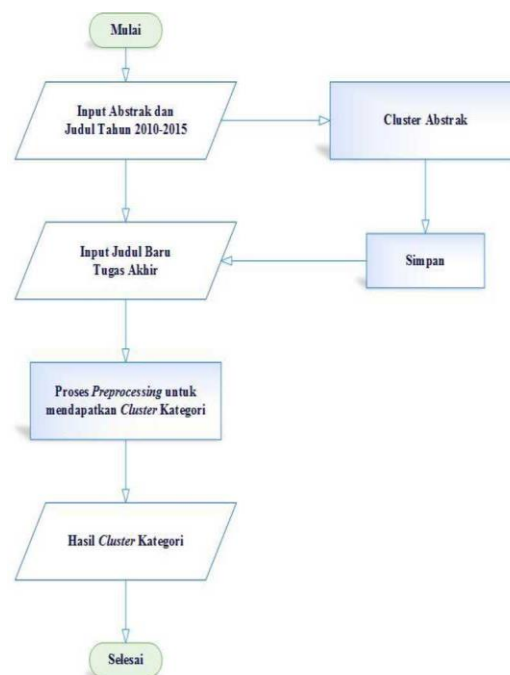
Oleh karena itu, pencarian informasi berdasarkan keyword/inputan ini sangat berguna untuk pencarian terarah dan berguna ketika user ingin mengetahui informasi yang berhubungan dengan *keyword* yang dicari (Adek, 2013). Begitu juga dengan pencarian informasi saat mahasiswa ingin mengetahui judul-judul tugas akhir yang mendekati yang akan diajukan. Karena setiap mahasiswa diharuskan untuk mengajukan judul tugas akhir, dimana judul tersebut harus merupakan judul yang baru dan belum pernah ada sebelumnya. Untuk itu dibutuhkan suatu sistem yang dapat memberikan informasi kepada mahasiswa yang ingin melakukan pencarian informasi untuk mengetahui informasi judul-judul tugas akhir yang mendekati ketika diajukan sehingga diharapkan dapat memperoleh hasil akhir informasi yang benar-benar relevan atau sesuai dengan inputan yang dicari.

Dalam penelitian ini akan dibahas tentang bagaimana cara untuk mengetahui tema judul tugas akhir yang akan diajukan yang memiliki kemiripan dan mendekati dengan judul tugas akhir yang sudah ada. Tugas akhir tersebut berdasarkan dari judul dan abstrak yang sudah ada kemudian dilakukan proses *clustering*. Selain judul, abstrak juga digunakan dalam proses *clustering* dikarenakan abstrak biasanya memberi gambaran secara ringkas dan jelas tentang isi suatu karya ilmiah (Puspitarani, 2013). Abstrak merupakan uraian singkat dan lengkap yang memuat judul permasalahan, pendekatan terhadap masalah, landasan teoritik yang digunakan, hasil temuan, dan rekomendasi. Metode yang digunakan pada proses pengklasteran dalam penelitian ini adalah metode *Single Linkage Hierarchical* (SLH). *Single Linkage* adalah proses pengklasteran yang didasarkan pada jarak terdekat antar obyeknya (*minimum distance*) (Devika, Sunitha, & Ganesh, 2016). Metode SLH sangat bagus untuk melakukan analisa pada tiap tahap pembentukan cluster (Handoyo, Mangkudjaja, & Nasution, 2014). Input untuk algoritma *single linkage* bisa berwujud jarak atau *similarities* antara

pasanganpasangan dari objek-objek. Kelompok-kelompok dibentuk dari *entities* individu dengan menggabungkan jarak paling pendek atau *similarities* (kemiripan) yang paling besar(Nuseir, Al-Ayyoub, Al-Kabi, Kanaan, & Al-Shalabi, 2017). Hasilnya akan ditampilkan judul-judul tugas akhir yang paling berhubungan dan memiliki kemiripan berdasarkan tema.

## METODE

Teknik pengumpulan data yang dilakukan pada tahap ini adalah dengan mengambil judul tugas akhir beserta abstrak pada tugas akhir mahasiswa pada jurusan Teknik Informatika Universitas Malikussaleh tahun 2010 - 2015. Tahap metode analisis ini merupakan tahap menganalisa sistem yang akan dibangun. Setelah analisa didapatkan maka langkah selanjutnya adalah membuat sebuah hasil analisa. Hasil analisa tersebut akan menjadi acuan dari perancangan sistem yang akan dibangun. Sistem yang akan dibangun dapat digambarkan pada gambar berikut:



Gambar 1 Skema sistem

Sistem yang akan dirancang adalah sebuah sistem pencarian yang mengambil objek tentang tugas akhir yang dikhususkan pada abstrak dan judul tugas akhir yang sudah ada. Bentuk aplikasi dirancang dalam bentuk menampilkan judul-judul tugas akhir mahasiswa yang memiliki kemiripan tema dengan *keyword* yang diinputkan oleh *user*. Judul yang ditampilkan pada proses pencarian merupakan judul-judul yang sudah dikelompokkan berdasarkan kategori pada proses clustering abstrak yang disesuaikan dengan judul dari abstrak tersebut.

Adapun bahasan tentang objek yang dipaparkan dalam *sistem* adalah judul dari tugas akhir yang sudah ada,

dimana judul-judul tersebut sudah disimpan didalam *database*. Begitu juga dengan abstrak, akan tetapi abstrak tidak ditampilkan pada *output user* melainkan hanya judul saja.

Perancangan system yang dibangun menggunakan pendekatan berorientasi objek. Berdasarkan kebutuhan yang akan diterapkan pada sistem ini, terdapat dua jenis kebutuhan yaitu kebutuhan sebagai *admin* dan *user*.

#### 1. Kebutuhan *admin*

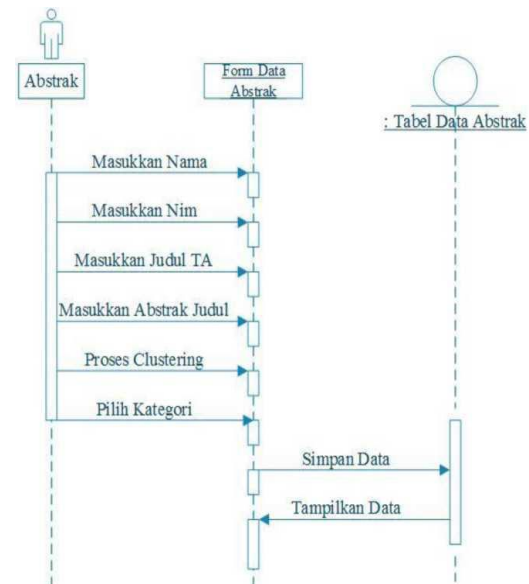
- Login sebagai *admin*
- Menambahkan data pada *admin*, *user*, judul dan abstrak kedalam *database*
- Mengubah data pada *admin*, *user*, judul dan abstrak yang terdapat didalam *database*
- Mengubah data pada *admin*, *user*, judul dan abstrak yang terdapat didalam *database*
- Melihat data *admin*, *user*, judul dan abstrak

#### 2. Kebutuhan *user*

- Login sebagai *user*
- Menginput tema judul tugas akhir untuk dilakukan proses pencarian
- Lihat hasil

## HASIL DAN PEMBAHASAN

Proses pelatihan dalam system yang dibangun dapat di gambarkan pada sequence diagram berikut:



Gambar 2 Sequence diagram pelatihan

Pada proses ini terdapat *field-field* yang harus dimasukkan dalam pengisian untuk menambah data abstrak, yaitu nama, Nim, judul Tugas Akhir, abstrak Judul, melakukan proses *clustering* kemudian memilih kategori, kemudian data tersebut disimpan kedalam *database* yang kemudian ditampilkan pada form data abstrak.

Dalam menyimpan data pelatihan, digunakan sebuah table untuk menyimpan informasi mengenai abstrak dari tugas akhir, dimana struktur dari table tersebut seperti pada table berikut:

Tabel 1 Tabel abstrak

Field name	Type	ket
Idabs	Int	ID dari judul
Nama	Varchar	Nama mahasiswa
Nim	Varchar	Nim mahasiswa
Judul_ta	Text	Judul tugas akhir
Abstrak_judul	Text	Abstrak dari tugas akhir
Kdkat	Varchar	Kategori dari abstrak judul

Selain dari abstrak, keyword dari tugas akhir juga disimpan kedalam database dan disimpan dalam table keyword berikut:

Tabel 2 Tabel keyword

Field name	Type	Ket
Id	Int	ID dari judul
Idabs	Int	ID dari abstrak judul
Kata	Varchar	Kata-kata dari abstrak
Frekuensi	Int	Banyaknya jmlah kata yang sama muncul
Bobot	Decimal	Perangkingan dari frekuensi

*Single Linkage* merupakan salah satu contoh dari bentuk pengelompokan data (*clustering*) dan biasa digunakan untuk menghitung tingkat kemiripan (Bhargava, Sharma, & Sharma, 2016; Hasdina & Rizal, 2016). Metode pengelompokan ini menggunakan objek yang paling dekat atau paling sama antar objek satu dengan yang lain untuk dikelompokkan. Metode ini menggunakan prinsip jarak minimum yang diawali dengan mencari

dua obyek terdekat dan keduanya membentuk cluster yang pertama (Devika et al., 2016). Pada langkah selanjutnya terdapat dua kemungkinan, diantaranya adalah obyek ketiga akan bergabung dengan cluster yang telah terbentuk, atau dua obyek lainnya akan membentuk cluster baru. Proses ini akan berlanjut sampai akhirnya terbentuk cluster tunggal. Pada metode ini jarak antar cluster didefinisikan sebagai jarak terdekat antar anggotanya.

Implementasi *Single Linkage* dilakukan dengan menghitung tingkat kemiripan antar cluster untuk jarak terdekat pada setiap cluster. Dari hasil penentuan tingkat kemiripan antar cluster tersebut dilakukan penggabungan antar cluster hingga terbentuk sebuah cluster tunggal. Pada penelitian ini, dalam pencarian kemiripan judul tugas akhir mahasiswa dilakukan dengan melakukan pengelompokan atau *clustering* berdasarkan kategori pada abstrak yang sudah ada didalam *database* terlebih dahulu untuk mendapatkan *cluster* baru untuk menentukan kategorinya berdasarkan abstrak. Kemudian setiap kata-kata yang terdapat pada abstrak di cari frekuensinya agar didapatkan nilai untuk sebuah matriks sehingga

perhitungan berdasarkan jarak terdekat dapat dilakukan.

Berikut adalah daftar kategori pada *cluster* yang terdapat pada system untuk menentukan kategori data baru berdasarkan abstrak menggunakan metode *Single Linkage*.

Tabel 3 Tabel kategori tugas akhir

kdkat (Kode Kategori)	nmkat (Nama Kategori)
K001	Jaringan
K002	Multimedia
K003	Pemograman
K004	Pengenalan Pola
K005	Pengolahan Citra
K006	Data Mining

Dalam sistem yang dibangun akan diterapkan metode *Single Linkage* sebagai formula pengelompokan kategori data berdasarkan abstrak yang diinputkan. Mengacu pada metode tersebut, maka proses pengelompokan kategori data berdasarkan abstrak yang diinputkan menggunakan matriks pencarian jarak terkecil. Proses pengelompokan kategori data berdasarkan abstrak yang diinputkan dimulai dengan menginputkan abstrak baru sebagai data uji, kemudian dapat di klik *button/tombol* 'Proses *Clustering*' maka system akan memproses secara otomatis dengan menggunakan metode

*Single Linkage* dan mencari kategori apakah data abstrak uji yang diinputkan dengan mengacu pada data training/latih yang terdapat pada *database*.

Dari data abstrak pada salah satu contoh data uji ketika dilakukan proses *clustering* dengan *persentase keyword* sebanyak 50%, didapatkan hasil abstrak termasuk kedalam kategori "Pemograman". Hasil abstrak uji didapat dengan beberapa proses yang akan diuraikan sebagai berikut. Abstrak uji diatas telah melalui proses *tokenizing* didapatkan kata dengan frekuensi sebagai berikut:

Tabel 4 Tabel tbToken\_uji

ID	Idabs	Kata	Frek
9290	1	Hadist	7
9291	1	Aplikasi	4
9489	1	Sistem	3
9490	1	Perkembangan	3
9491	1	Perangkat	2
9493	1	Diagram	2
9495	1	Mempermudah	2
9496	1	Metode	2

Setelah melalui proses *tokenizing*, pada system berdasarkan jumlah presentase keyword sebanyak 50%, maka didapatkan keyword uji sebagai berikut:

Tabel 5 tabel TbKeyword\_Uji

Id	Idabs	Kata	Frek	Bobot
6130	1	Hadist	7	36.88
6131	1	Aplikasi	4	21.07
6132	1	System	3	15.80
6133	1	Perkembangan	3	15.80
6134	1	Perangkat	2	10.54

Dimana perhitungan bobot didapatkan berdasarkan:

$$W = tf * \log \left( \frac{n}{N} \right) \quad (1)$$

Dimana

- W : bobot terhadap kata,
- Tf : term frekuensi (banyaknya kata dalam dokumen)
- n : total dokumen
- N : banyaknya dokumen yang mengandung kata yang dicari.

Kemudian, berdasarkan abstrak uji diatas setelah melalui proses *tokenizing* dan hasil *keyworduji* tersebut nilai matriks dan proses metode *single linkage* sebagai berikut:

- Iterasi pertama

	1	2	3	4	5
1	0.00				
2	3.00	0.00			
3	2.00	5.00	0.00		
4	<b>1.00</b>	4.00	1.00	0.00	
5	1.00	4.00	1.00	0.00	0.00

Nilai jarak terkecil pada iterasi pertama adalah 1.00 yang terdapat pada baris ke 4 dan kolom ke 1, maka baris 4 dan kolom 1 bergabung, sehingga

membentuk matriks baru pada iterasi kedua.

- Iterasi kedua

	14	2	3	5
14	0.00			
2	3.00	0.00		
3	2.00	5.00	0.00	
5	<b>1.00</b>	4.00	1.00	0.00

Nilai jarak terkecil pada iterasi 2 adalah = 1,00 yang terdapat pada baris ke 5 dan kolom 14, maka baris ke 5 dan kolom ke 14 bergabung sehingga membentuk matriks baru pada iterasi 3:

- Iterasi ketiga

	145	2	3
145	0.00		
2	3.00	0.00	
3	<b>2.00</b>	5.00	0.00

Nilai jarak terkecil pada iterasi 3 adalah = 2,00 yang terdapat pada baris3 dan kolom 145, maka baris 3 dan kolom 145 bergabung sehingga membentuk matriks baru pada iterasi 4:

- Iterasi keempat

	1345	2
1345	0.00	
2	3.00	0

Nilai jarak terkecil pada iterasi terakhir adalah 3.00.

Setelah mendapatkan nilai iterasi terakhir pada proses perbandingan frekuensi dengan menggunakan proses metode *single linkage*, kemudian,

dilakukan proses pencarian kemiripan kata-kata yang bersesuaian antara tabel *keyword uji* dan tabel *keyword* yang merupakan data training untuk mengetahui frekuensi berdasarkan *abstrak\_keyword* yang bertujuan mendapatkan nilai total frekuensi pada kategori. Dengan demikian didapatkan hasil total frekuensi pada masing-masing kategori adalah sebagai berikut:

Tabel 6 Total frekuensi kategori

No	Kode	Frekuensi	Total
1	K003	Pemograman	53
2	K002	Multimedia	32
3	K006	Data Mining	10
4	K001	Jaringan	6
5	K004	Pengenalan	5
6	K005	Pengolahan	2

Sebelumnya, didapatkan nilai jarakterkecil pada iterasi terakhir yaitu 3,00. Kemudian hasil cluster dari iterasiterakhir dibandingkan dengan hasilcluster datatraining abstrak berdasarkan total frekuensi setiap kategori untuk mengetahui jarak cluster pada tiap-tiap kategori. Dalam penelitian ini menghitung nilai jarakcluster dengan Euclidean distance:

$$D_{ij} = \sqrt{\sum_{k=1}^n (d_{jk} - c_{jk})^2} \quad (2)$$

Dimana

D merupakan jumlah atribut

$D_{jk}$  dan  $C_{jk}$  merupakan data yang akan dihitung jaraknya.

Hasil nilai diatas membantu kita mengetahui jarak cluster pada setiap kategori dengan data training yang diinputkan, dimana dapat dilihat pada tabel berikut:

Tabel 7 Jarak cluster

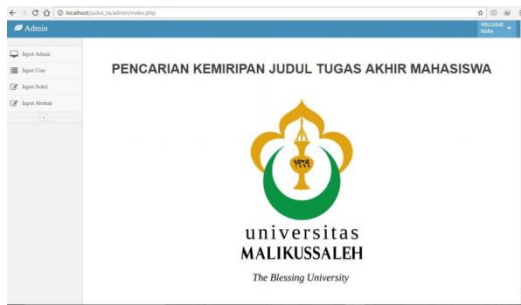
No	Kode	Kategori	Frekuensi	Jarak
1	K003	Pemograman	53	50
2	K002	Multimedia	21	18
3	K005	Pengolahan Citra	10	7
4	K004	Pengenalan Pola	6	3
5	K006	Data Mining	5	2
6	K001	Jaringan	2	1

Dengan demikian, disimpulkan hasil berdasarkan nilai tertinggi dari total frekuensi adalah kategori Pemograman (K003) dengan jarak cluster 50, oleh karena itu abstrak uji tersebut digolongkan kepada kategori Pemograman.

### Interface Sistem

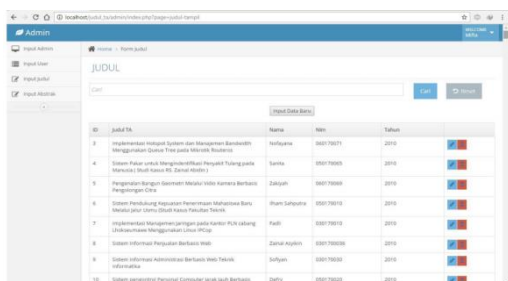
Untuk mengimplementasikan sistem, aplikasi ini dibuat menggunakan bahasa pemrograman PHP dengan bantuan XAMPP yaitu MySQL sebagai tempat penampungan data pada system. Pada system ini terdapat 2 aktor yaitu *admin* dan *user* dengan tampilan dan fungsi yang berbeda-beda.





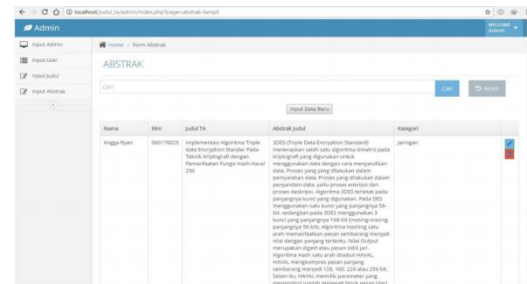
Gambar 3 Menu utama aktor admin

Setelah *admin* berhasil melakukan *login* maka *system* akan menampilkan menu *Home* pada halaman *admin*. *Home* merupakan menu utama pada halaman *admin*, di sebelah kiri *dashboard* menu *Home* tersedia empat buah *button* yang gunanya untuk mengakses menu-menu yang ada pada halaman *admin*. *Button* “Input Admin” berguna untuk menampilkan dan menambah data baru admin, *button* “Input User” berguna untuk menampilkan dan menambah data baru user, *button* “Input Judul” berguna untuk menampilkan dan menambah data baru judul, *button* “Input Abstrak” berguna untuk menampilkan dan menambah data baru abstrak.



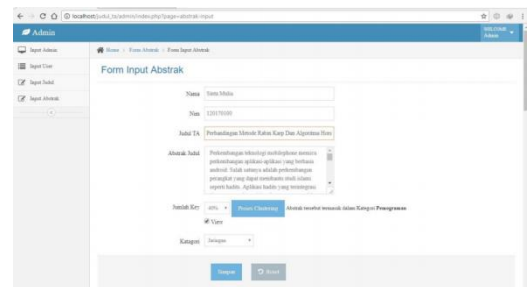
Gambar 4 Menu data judul

Gambar diatas menunjukkan tampilan data judul-judul tugas akhir mahasiswa teknik informatika Universitas Malikussaleh dari tahun 2010 sampai dengan tahun 2015 pada halaman judul.



Gambar 5 Menu data abstrak

Gambar 5 menunjukkan tampilan data abstrak berdasarkan judul-judul tugas akhir mahasiswa dari tahun 2010 sampai dengan tahun 2015 pada halaman abstrak.



Gambar 6 Form input data abstrak

Gambar diatas menunjukkan tampilan input abstrak untuk menambah data abstrak yang akan ditampilkan pada menu abstrak dan menentukan kategori pada halaman abstrak.



Gambar 7 Tampilan untuk *user* pengguna

Setelah *user* berhasil melakukan *login* maka *system* akan menampilkan menu *Home* pada halaman *user*. *Home* merupakan menu utama pada halaman *user*, pada tampilan *dashboard* menu *Home* dapat dilihat hanya ada form untuk menginput tema judul Tugas Akhir bagi *user*, kemudian setelah *user* menginput tema judul tugas akhir yang ingin dicari kemiripannya dengan judul yang sudah ada dengan judul yang diinputkan oleh *user* tersebut. Kemudian *user* dapat menekan tombol “cari” untuk mendapatkan hasil pencarian kemiripan judul sesuai dengan tema yang *user* inputkan. Disini *user* hanya diberi hak akses untuk mencari kemiripan tema judul tugas akhir dan mendapatkan hasilnya.

Kategori	Judul	Bobot	Nim	Nama
Data Mining	Aplikasi Penemuan Status GDA Menggunakan Algoritma Naive Bayes	47,32	10017008	Rydeli Akmal
Data Mining	Implementasi Metode Rough Set Untuk Memeriksa Data Hasil Penelitian Mendapatkan Program Pada PT. BANK BUKIT BARIS (BANK BUKIT BARIS) LEBANG	22,73	10017009	Jahara
Data Mining	ALGORITMA DATA MINING KLASIFIKASI MENGGUNAKAN ALGORITMA NAIVE BAYES	22,18	98017003	Melisa

Gambar 8 Hasil pencarian judul

Pada menu *user* terdapat kotak tempat menginputkan tema judul untuk dicari kemiripannya oleh *user* berdasarkan *keyword* dan bobot pada abstrak untuk perankingan, bobot terbesar akan muncul pada baris teratas. Pada tampilan hasil terdapat keterangan kategori, output judul tugas akhir yang memiliki hubungan dengan inputan, bobot, Nim, dan nama mahasiswa.

Pada sistem pencarian kemiripan judul tugas akhir mahasiswa menggunakan metode *Single Linkage Hierarchical*, metode *Single Linkage Hierarchical* digunakan pada halaman *admin* ketika dilakukan proses pengkategorian data. Data yang diinputkan dicari kecocokan dengan dengan abstrak training yang terdapat dalam database, kemudian *system* memberikan *output* berupa kategori yang memiliki kecocokan dengan yang data uji yang diinputkan sesuai dengan *keyword*.

Pada saat pengujian dengan 60 abstrak sebagai data uji, tingkat keberhasilan pada kategori Multimedia, Pemrograman dan Pengolahan Citra mencapai 96,77%. Sedangkan pada kategori Jaringan, Pengenalan Pola dan Data Mining persentase keberhasilannya sebesar 82.2%. Dimana persentase

keberhasilan kecocokan pada kategori multimedia sebesar 100%, kategori pemograman sebesar 100%, kategori pengolahan citra sebesar 87,5%, kategori pengenalan pola sebesar 11,11%.

Untuk mendapatkan hasil yang lebih optimal, pada pengembangan system ini selanjutnya dapat menggunakan metode *clustering* yang lainnya, karena penggunaan metode *Single Linkage Hierarchical* pada sistem pencarian kemiripan judul tugas akhir mahasiswa saat proses pencocokan kategori antara data uji dan data training masih terjadi kekeliruan hasil dan pembentukan clusternya kurang optimal, dikarenakan metode ini hanya mengenali jarak tanpa memperhatikan makna yang terkandung pada tiap-tiap atribut.

## KESIMPULAN DAN SARAN

### 1. Kesimpulan

Berdasarkan hasil penelitian diatas dapat diambil kesimpulan sebagai berikut :

- a. Hasil dari percobaan 60 data uji abstrak persentase keberhasilan kecocokan pada kategori multimedia sebesar 100%

- b. Pengujian Kategori pemograman sebesar 100%.
- c. Pengujian kategori pengolahan citra sebesar 87,5%.
- d. Pengujian kategori pengenalan pola sebesar 11,11%.
- e. Pengujian pada kategori jaringan dan data mining tidak ditemukan kecocokan. Dan pada halaman *user*, hasil berupa judul tugas akhir yang ada pada *database* berdasarkan dengan tema judul tugas akhir yang di-*input*-kan oleh *user*.

### 2. Saran

Saran yang dapat diambil untuk pengembangan lebih lanjut adalah sebagai berikut :

- a. Dapat menggunakan metode kemiripan yang lain dalam pengujian kasus tugas akhir mahasiswa.
- b. Metode *Single Linkage Hierarchical* dalam kasus yang lain dalam penentuan kemiripan.

## DAFTAR PUSTAKA

- Adek, R. T. 2013. Sistem Pendukung Keputusan Seleksi Calon Penerima Beasiswa Pada Universitas Malikussaleh. *TECHSI*, 5(1), 113–124.  
<https://doi.org/10.29103/Techsi.V5i1.144>

- Aggarwal, C. C., Rendy, C. K. 2013. *Data Clustering Algorithms And Applications*. Minneapolis, Minnesota: Taylor & Francis Group.
- Bhargava, R., Sharma, Y., Sharma, G. (2016). ATSSI: Abstractive Text Summarization Using Sentiment Infusion. *Procedia Computer Science*, 89, 404–411. <https://doi.org/10.1016/j.procs.2016.06.088>
- Devika, M. D., Sunitha, C., Ganesh, A. (2016). Sentiment Analysis: A Comparative Study On Different Approaches. *Procedia Computer Science*, 87, 44–49. <https://doi.org/10.1016/j.procs.2016.05.124>
- Hachaj, T., Ogiela, M. R. 2017. Clustering Of Trending Topics In Microblogging Posts: A Graph-Based Approach. *Future Generation Computer Systems*, 67, 297–304. <https://doi.org/10.1016/j.future.2016.04.009>
- Handoyo, R., Mangkudjaja, R., Nasution, S. M. 2014. Perbandingan Metode Clustering Menggunakan Metode Single Linkage Dan K-Means Pada Pengelompokan Dokumen. *JSM STMIK Mikroskil*, 15(2), 73–82.
- Hasdina, N., Rizal, R. 2016. IMPLEMENTASI METODE CUSUM (CUMMULATIVE SUMMARY) UNTUK MENENTUKAN DAERAH RAWAN KECELAKAAN BERBASIS WEB DI KOTA LHOKSEUMAWE. *TECHSI-Jurnal Teknik Informatika*, 8(1).
- Nuseir, A., Al-Ayyoub, M., Al-Kabi, M., Kanaan, G., & Al-Shalabi, R. (2017). Improved hierarchical classifiers for multi-way sentiment analysis. *International Arab Journal of Information Technology*, 14(4A Special Issue), 654–661.
- Puspitarani, Y. 2013. Pemanfaatan Clustering dalam Pencarian Kemiripan Dokumen Paper Conference (pp. 21–26).