**JURNAL**

# PENERAPAN *COMPLETE AND AVERAGE LINKAGE* PADA PEMBENTUKAN *RESEARCH GROUP* DOSEN

## *IMPLEMENTATION OF COMPLETE AND AVERAGE LINKAGE TO THE FORMATION FACULTY RESEARCH GROUP*


**Oleh:**

**DEVY SURYANINGTYAS**

**12.1.03.02.0366**

**Dibimbing oleh :**

1. Irwan Setyo Widodo, M.Si
2. Daniel Swanjaya.,M.Kom

**PROGRAM STUDI TEKNIK INFORMATIKA**

**FAKULTAS TEKNIK**

**UNIVERSITAS NUSANTARA PGRI KEDIRI**


**2017**

# SURATPERNYATAAN
# ARTIKEL SKRIPSI TAHUN 2017

**Yang bertanda tangan dibawah ini:**

| | |
|---|---|
| Nama Lengkap | :DEVY SURYANINGTYAS |
| NPM | :12.1.03.02.0366 |
| Telepon/HP | :085736925751 |
| Alamat Surel (Email) | :devysuryaningtyas@gmail.com |
| Judul Artikel | :PENERAPAN COMPLETE AND AVERAGE LINKAGE PADA PEMBENTUKAN RESEARCH GROUP DOSEN |
| Fakultas – Program Studi | :FT-TEKNIK INFORMATIKA |
| Nama Perguruan Tinggi | :UNIVERSITAS NUSANTARA PGRI KEDIRI |
| Alamat Perguruan Tinggi | :JL. K.H. Achmad Dahlan No. 76 Tel. : (0354) 771576, 771503, 771495 Kediri |

Dengan ini menyatakan bahwa:

a. artikel yang saya tulis merupakan karya saya pribadi (bersama tim penulis) dan bebas plagiarisme;

b. artikel telah diteliti dan disetujui untuk diterbitkan olehDosen Pembimbing I dan II.

Demikian surat pernyataan ini saya buat dengan sesungguhnya. Apabila dikemudian hari ditemukan ketidaksesuaian data dengan pernyataan ini dan atau ada tuntutan dari pihak lain, saya bersedia bertanggung jawab dan diproses sesuai dengan ketentuan yang berlaku.

| Mengetahui | | Kediri, 30 Januari 2017 |
|---|---|---|
| Pembimbing I | Pembimbing II | Penulis, |
| Irwan Setyo Widodo, M.Si<br>NIDN 0701098404 | Daniel Swanjaya, M.Kom<br>NIDN 0723098303 | Devy Suryaningtyas<br>12.1.03.02.0366 |



Devy Suryaningtyas | 12.1.03.02.0366
FT – Teknik Informatika
simki.unpkediri.ac.id
|| 1||

# PENERAPAN *COMPLETE AND AVERAGE LINKAGE* PADA PEMBENTUKAN *RESEARCH GROUP* DOSEN

Devy Suryaningtyas
12.1.03.02.0366
Teknik – Teknik Informatika
devysuryaningtyas@gmail.com
Irwan Setyo Widodo, M.Si dan Daniel Swanjaya.,M.Kom
UNIVERSITAS NUSANTARA PGRI KEDIRI

## ABSTRAK

**Devy Suryaningtyas** : Penerapan Complete And Average Linkage Pada Pembentukan Research Group Dosen, Skripsi, Teknik Informatika, Fakultas Teknik, UN PGRI Kediri, 2017.

Penelitian ini dilatar belakangi oleh Dosen – dosen yang memiliki kesamaan/kemiripan keahlian dan sering berkolaborasi membentuk kelompok penelitian atau grup *riset* dalam lingkup jurusan. Akan tetapi di UN PGRI Kediri khususnya prodi Teknik Informatika dan Sistem Informasi belum ada kelompok penelitian atau grup *riset.*

Permasalahan penelitian ini adalah bagaimana mengimplementasikan *Hierarchical Clustering* pada pengelompokan dosen pembimbing skripsi?. Metode untuk pembentukan *research group* dosen berdasarkan data skripsi mahasiswa menggunakan *Hierarchical Clustering* yaitu *Complete and Average Linkage*, sehingga memungkinkan untuk membentuk *grup riset* agar para dosen membimbing skripsi sesuai dengan bidang penelitiannya masing- masing.

Hasil dari penelitian penerapan sistem dengan metode *Hierarchical Clustering* cukup baik. Untuk penentuanya dicari hasil nilai *cosine* yang terbesar yaitu *complete linkage* kemudian dikelompokkan sesuai nilai yang dicari. Hasil pengelompokkan berupa kode dosen serta data skripsi yang sesuai dengan bidang penilitian dosen.

**KATA KUNCI :** *Research Group* Dosen , *Hierarchical Clustering, Complete and Average Linkage*.

## I. LATAR BELAKANG

Dosen pembimbing skripsi merupakan salah satu faktor pendukung dalam penyelesaian penyusunan proposal skripsi. Sebagai bagian dari proses penulisan skripsi, peran dosen sebagai pembimbing skripsi menjadi sangat penting karena merupakan tanggungjawab dosen untuk memastikan bahwa mahasiswa mampu menyusun skripsi dengan baik hingga skripsi siap diajukan dan berkualitas.

Dalam penulisan sebuah karya tulis (skripsi) tentunya melibatkan beberapa dosen pembimbing, kerjasama ini sering disebut dengan kolaborasi. Kolaborasi yang terjalin ini memiliki konsistensi, dimana masing-masing memiliki pasangan yang tetap sesuai dengan bidang penelitian yang

diminati. Dosen – dosen yang memiliki kesamaan/kemiripan keahlian dan sering berkolaborasi ini membentuk kelompok penelitian atau grup *riset* dalam lingkup jurusan. Tetapi di UN PGRI Kediri khususnya prodi Teknik Informatika dan Sistem Informasi belum ada kelompok penelitian atau grup *riset.* Oleh karena itu, perlu adanya sistem yang dapat digunakan sebagai pengelompokan penelitian atau grup *riset*. Dengan menggunakan metode *Hierarchical Clustering* dapat membantu pengelompokan sistem tersebut untuk mempermudah pekerjaan. Metode *Hierarchical Clustering* adalah suatu algoritma *clustering* yang dapat digunakan untuk meng-*cluster* dokumen *(document clustering*). Dari teknik *hierarchical clustering*, dapat dihasilkan suatu kumpulan partisi yang berurutan, dimana dalam kumpulan tersebut terdapat *cluster-cluster* yang mempunyai poin-poin individu, kemudian yaitu *cluster* yang didalamnya terdapat poin-poin yang dipunyai semua *cluster* didalamnya.

Dengan metode tersebut, diharapkan pengelompokan data skripsi berdasarkan dosen pembimbing akan lebih akurat, karena seharusnya yang akan dibuat didasarkan pada nilai kriteria dan bobot yang sudah ditentukan sehingga akan mendapatkan hasil yang akurat. Dari uraian di atas penulis mencoba mengangkat masalah tersebut dengan judul "**Penerapan *Complete* and *Avarage Linkage* pada Pembentukan *Research Group* Dosen** ". Berdasarkan permasalahan tersebut, maka perlu dibuat suatu sistem untuk memberikan pertimbangan yang dapat membantu tugas prodi Teknik Informatika dan Sistem Informasi UN PGRI Kediri.

## II. METODE

### 1. Cosine Similarity

Perhitungan *cosine similarity* yang memperhitungkan perhitungan pembobotan kata pada suatu dokumen dapat dinyatakan dengan perumusan :

$$CosSim(d_i, q_i) = \frac{q_i \bullet d_i}{|q_i||d_i|} = \frac{\sum_{j=1}^{t}(q_{ij}d_{ij})}{\sqrt{\sum_{j=1}^{t}(q_{ij})^2 \sum_{j=1}^{t}(d_{ij})^2}}$$

Keterangan :

*qij* = bobot istilah *j* pada dokumen *i* = *tfij* **idfj*

*dij* = bobot istilah *j* pada dokumen *i* = *tfij* * *idfj*

### 2. Stemming Algoritma Tala

Proses *stemming* adalah proses untuk mencari *root* dari kata hasil dari proses *filtering*.

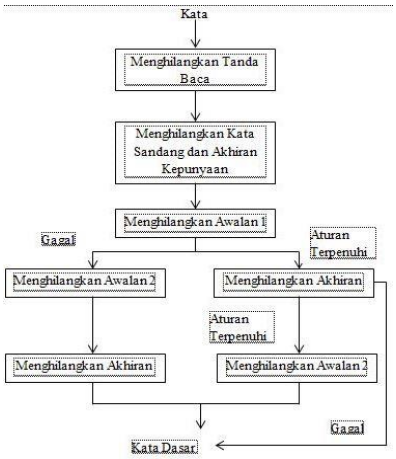
Pencarian *root* sebuah kata atau biasa disebut degan kata dasar dapat memperkecil hasil indeks tanpa harus menghilangkan makna. *Filtering* adalah proses pengambilan kata-kata yang dianggap penting atau mempunyai makna. Ada dua pendekatan pada proses *stemming* yaitu pendekatan kamus dan pendekatan aturan. Beberapa penelitian juga telah dilakukan untuk stemmer bahasa Indonesia baik untuk pendekatan kamus ataupun pendekatan aturan. Ahmad,Vega, Jelita dan Tala mereka masing-masing mempunyai algoritma yang berbeda dalam melakukan proses *stemmer* pada dokumen bernahasa Indonesia.

*Stemmer* tala merupakan adopsi dari algoritma *stemmer* bahasa inggris terkenal *porter stemmer*. *Stemmer* ini menggunakan *rule base* analisis untuk mencari *root* sebuah kata. *Stemmer* ini sama sekali tidak menggunakan kamus sebagai acuan, seperti halnya *stemmer* ahmad,vega dan jelita.

Pada *stemmer* Tala terdapat 5 langkah utama dengan 3 langkah awal dan 2 langkah pilihan, langkah-langkah tersebut sbb:

1. Menghilangkan partikel
2. Menghilangkan kata sandang dan kepunyaan.
3. Menghilangkan awalan 1
4. Jika suatu aturan terpenuhi jalankan sbb :
   a. Hilangkan Akhiran
   b. Jika suatu aturan terpenuhi, hilangkan awalan 2. Jika tidak proses *stemming* selesai
5. Jika tidak ada aturan yang terpenuhi jalankan sbb :
   a. Hilangkan awalan 2.
   b. Hilangkan Akhiran
   c. Proses *stemming* selesai.

Selain itu, tala membagi imbuhan menjadi 5 *cluster* / kelompok. Alur proses dari algoritma Tala diperlihatkan pada gambar 2.1 berikut :



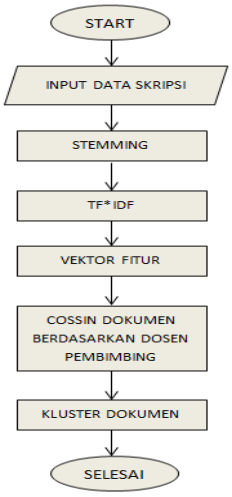

**Gambar 2.1** Proses *Stemming* Algoritma Tala

# III. HASIL DAN KESIMPULAN

## A. HASIL

### 1. Desain Sistem

Sistem yang telah dirancang digunakan untuk penerapan *complete and average linkage* pada pembentukan *research group* dosen. Proses diawali dengan melakukan *input* data skripsi, setelah itu akan masuk proses *stemming*, setelah melakukan proses *stemming* akan dilakukan proses *tf*idf* yang kemudian hasilnya akan di *vektor* fitur, selanjutnya yaitu proses *cossim* dokumen berdasarkan dosen pembimbing kemudian dokumen akan dikuluster. Proses dapat dilihat pada flowchart di bawah ini :

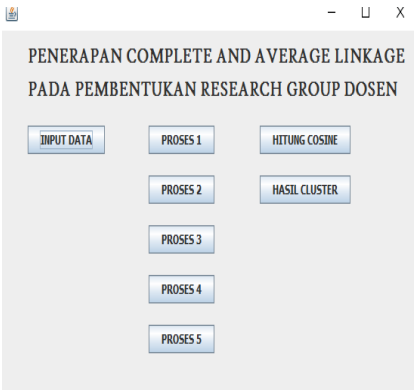



**Gambar 3.1** *Flowchart* Sistem

Dari diagram di atas didapatkan deskripsi seperti di bawah ini :

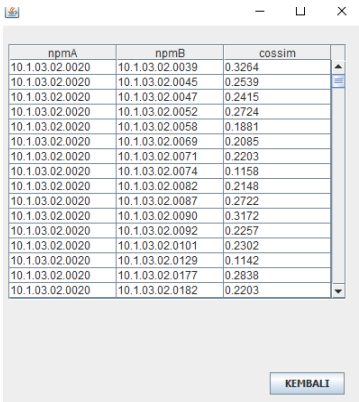
1. *Input* data skripsi.
2. *Stemming* digunakan untuk mengembalikan kata ke bentuk semula.
3. *TF*IDF* digunakan untuk menghitung banyaknya kata yang sering muncul pada dokumen.
4. *Vektor fitur* digunakan untuk menghitung banyaknya fitur dokumen.
5. *Cosine* dokumen digunakan untuk perhitungan pembobotan kata pada dokumen.
6. Kluster dokumen digunakan untuk pencarian hasil pengelompokan dosen.

### 2. Tampilan Menu Utama




**Gambar 3.2** Tampilan Menu Utama

### 3. Tampilan Hasil *Cosine*



| npmA | npmB | cossim |
|---|---|---|
| 10.1.03.02.0020 | 10.1.03.02.0039 | 0.3264 |
| 10.1.03.02.0020 | 10.1.03.02.0045 | 0.2539 |
| 10.1.03.02.0020 | 10.1.03.02.0047 | 0.2415 |
| 10.1.03.02.0020 | 10.1.03.02.0052 | 0.2724 |
| 10.1.03.02.0020 | 10.1.03.02.0058 | 0.1881 |
| 10.1.03.02.0020 | 10.1.03.02.0069 | 0.2085 |
| 10.1.03.02.0020 | 10.1.03.02.0071 | 0.2203 |
| 10.1.03.02.0020 | 10.1.03.02.0074 | 0.1158 |
| 10.1.03.02.0020 | 10.1.03.02.0082 | 0.2148 |
| 10.1.03.02.0020 | 10.1.03.02.0087 | 0.2722 |
| 10.1.03.02.0020 | 10.1.03.02.0090 | 0.3172 |
| 10.1.03.02.0020 | 10.1.03.02.0092 | 0.2257 |
| 10.1.03.02.0020 | 10.1.03.02.0101 | 0.2302 |
| 10.1.03.02.0020 | 10.1.03.02.0129 | 0.1142 |
| 10.1.03.02.0020 | 10.1.03.02.0177 | 0.2838 |
| 10.1.03.02.0020 | 10.1.03.02.0182 | 0.2203 |



**Gambar 5.5** Tampilan Hasil *Cosine*

## B. KESIMPULAN

Setelah melakukan analisis, perancangan, implementasi dan pengujian pada bab sebelumnya, maka dapat disimpulkan bahwa aplikasi telah berhasil dibuat dan diujicobakan untuk mengimplementasikan *Hierarchical Clustering* pada pengelompokan dosen pembimbing skripsi.

## IV. DAFTAR PUSTAKA

Alfina, Tahta, Budi Santosa dan Ali Ridho Barakbah. 2012. *Analisa Perbandingan Metode Hierarchical Clustering, K-means dan Gabungan Keduanya dalam Cluster Data*. Surabaya : Institut Teknologi Sepuluh Nopember (ITS).

Handoyo, Rendy, R. Rumani M dan Surya Michrandi Nasution. 2014. *Perbandingan Metode Clustering Menggunakan Metode Single Linkage dan K - Means pada Pengolompokan Dokumen*. Bandung : Universitas Telkom.

Pradnyana, Gede Aditra dan Ngurah Agus Sanjaya ER. 2012. *Perancangan dan Implementasi Automated Document Integration dengan Menggunakan Algoritma Complete Linkage Agglomerative Hierarchical Clustering.* Bali : Universitas Udayana.

Swanjaya, Daniel, Chastine Fatichah dan Diana Purwitasari. 2015. *Hypergraph – Partitioning Pada Co – Authorship Graph Untuk Pengelompokan Penulis Berdasarkan Topik Penelitian*. Surabaya : Institut Teknologi Sepuluh Nopember (ITS).

Utomo, Mardi Siswo. 2013. *Implementasi Stemmer Tala pada Aplikasi Berbasis Web*. Program Studi Teknik Informatika, Universitas Stikubank.