

Penerapan Teknik Single Linkage Dalam Klasterisasi Mahasiswa Berdasarkan Indikator Pembidangan Skripsi

Eci Marcelina Silalahi*, Efori Bu'ulolo

Program Studi Teknik Informatika, Universitas Budi Darma, Medan, Indonesia

Email: ¹silalahiecmarcelina@gmail.com

Submitted: 09/06/2020; Accepted: 28/06/2021; Published: 30/06/2021

Abstrak—Program studi Informatika Komputer STMIK Pelita Nusantara Medan menyediakan beberapa bidang ilmu yang dapat dipilih mahasiswa dalam menyelesaikan tugas akhir yang berupa skripsi. Selama ini didalam pelaksanaannya, penentuan bidang skripsi ini dipilih sendiri oleh mahasiswa yang akan mengajukan judul skripsi sebagai salah satu syarat untuk memperoleh gelar sarjana di STMIK Pelita Nusantara Medan tanpa adanya arahan dari pihak kampus baik dari bagian akademik maupun ketua program studi. Oleh karena itu, tidak sedikit mahasiswa yang akhirnya memilih bidang ilmu yang tidak sesuai dengan kemampuan dirinya. Berdasarkan masalah yang ada mengenai pembidangan skripsi ini maka dibuatlah sistem klasterisasi (pengklasteran) mahasiswa berdasarkan indikator-indikator pembidangan skripsi dengan menerapkan salah satu metode data mining yaitu clustering dimana teknik yang digunakan adalah Teknik Single Linkage. Teknik Single Linkage merupakan cara untuk mengelompokkan objek atau individu berdasarkan ukuran kemiripan kriteria atau jarak terdekat antar objek (minimum distances). Dengan demikian, hasil dari pengklasteran (klasterisasi) yang dirancang dapat membantu pihak akademik STMIK Pelita Nusantara Medan khususnya Ketua Program Studi Informatika Komputer dalam mengarahkan mahasiswa dalam penentuan bidang skripsi sehingga mahasiswa memilih bidang skripsi yang sesuai dengan kemampuannya.

Kata Kunci: Data Mining; Single Linkage; Bidang; Skripsi

Abstract—Pelita Nusantara Medan STMIK Computer Informatics study program provides several fields of science that students can choose in completing their final thesis in the form of a thesis. During this time in its implementation, the determination of the field of this thesis was chosen by the student who will submit the thesis title as one of the requirements to obtain a bachelor's degree at STMIK Pelita Nusantara Medan without any direction from the campus either from the academic department or the head of the study program. Therefore, not a few students who ultimately choose a field of science that is not in accordance with their abilities. Based on the existing problems regarding the development of this thesis, a student clustering system was made based on the indicators of thesis development by applying one of the data mining methods, namely clustering, where the technique used was the Single Linkage Technique. Single Linkage Technique is a way to group objects or individuals based on criteria similarity size or closest distance between objects (minimum distances). Thus, the results of clustering (clustering) that are designed can help the academic STMIK Pelita Nusantara Medan, especially the Chairperson of the Computer Informatics Study Program in directing students in the determination of the field of thesis so that students choose the field of thesis in accordance with their abilities.

Keywords: Data Mining; Single Linkage; Field; Thesis

1. PENDAHULUAN

Sekolah Tinggi Manajemen Informatika dan Komputer (STMIK) Pelita Nusantara Medan merupakan salah satu Perguruan Tinggi Swasta yang memiliki 2 program studi, yaitu: S1 Teknik Informatika, D3 Manajemen Informatika. Sama halnya dengan Perguruan tinggi yang lain, STMIK Pelita Nusantara Medan juga mewajibkan seluruh mahasiswanya untuk menyelesaikan tugas akhir atau skripsi sebagai salah satu syarat mutlak untuk dapat lulus dan memperoleh gelar ahlimadya ataupun sarjana komputer. Dalam hal ini, program studi Teknik informatika menyediakan beberapa bidang atau disiplin ilmu dalam hal pembuatan skripsi mahasiswa. Adapun beberapa diantaranya, yaitu: Data Mining, Sistem Pendukung Keputusan (Decision Support System), Sistem Pakar (Expert System), Jaringan Syaraf Tiruan (Neural Network), Keamanan Komputer (Computer Security), Pemodelan dan Simulasi serta Pengolahan Citra (Image Processing)

Kesulitan yang dialami mahasiswa dalam menentukan judul skripsi, dimana skripsi yang memiliki arti sebagai persyaratan terakhir setiap mahasiswa untuk mencapai gelar sarjana (S1) pada setiap jurusan masing-masing. Disini mahasiswa yang memilih judul tidak berdasarkan minat dan nilai-nilai yang ditentukan dan sulit dipahami setiap mahasiswa. Sehingga banyak mahasiswa yang tidak menyelesaikan skripsi, tidak lulus tepat waktu karena kurang memahami arti dari judulnya tersebut.

Data Mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Dalam metode clustering ada beberapa teknik dan salah satu diantaranya adalah teknik Single Linkage. Konsep penyelesaian Teknik Single Linkage dalam hal ini adalah menghitung nilai kedekatan (proximity) terhadap data empiris yang berkenaan dengan pembidangan skripsi di STMIK Pelita Nusantara Medan. Dalam masalah yang dibahas pada penelitian ini akan dirancang sebuah perangkat lunak berbasis Dekstop Programming yang diharapkan menjadi solusi pemecahan masalah.

Dekstop Programming adalah suatu aplikasi pemrograman berbasis dekstop. Perangkat lunak yang akan dirancang mengadopsi teknik single linkage. Konsep perancangan akan dilakukan dengan cara menganalisis masalah dan kebutuhan dalam penelitian yang dibahas kemudian dilakukan sebuah rating terhadap variabel-

variabel atau indikator-indikator penyebab permasalahan dan pada fase akhir akan dilakukan sebuah perancangan sistemnya sehingga dapat menyelesaikan masalah sesuai dengan yang diharapkan.

Pada penelitian ini membahas tentang clustering data untuk data teks, yaitu menggunakan data berita yang diambil dari situs online dan dipindah ke notepad dengan format .txt. Dengan menggunakan metode single linkage clustering [1].

Penelitian ini menggunakan suatu program java untuk melakukan monitoring terhadap kinerja Agent hasil monitoring didapat lalu dilakukan pengelompokan menggunakan dua metode yaitu perhitungan manual dan metode Single Linkage [2].

Dari permasalahan penelitian ini, diharapkan perangkat lunak yang dirancang dapat membantu STMIK Pelita Nusantara Medan dalam hal ini mengarahkan pembidangan skripsi kepada mahasiswa yang selama ini belum ada acuan atau sistem yang digunakan dalam menentukan bidang skripsi yang sesuai dengan mahasiswa sehingga pemilihan bidang skripsi mahasiswa belum efektif karena tidak berdasarkan kemampuan melainkan hanya berdasarkan keinginan mahasiswa atau adanya pengaruh dari teman maupun faktor lain.

2. METODOLOGI PENELITIAN

2.1 Data Mining

Data mining adalah serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual. *Data mining* juga merupakan bidang dari beberapa keilmuan yang menyatukan teknik dari pembelajaran mesin, pengenalan pola, statistik, *database*, dan visualisasi untuk penanganan permasalahan pengambilan informasi dari *database* yang besar [1]. Dalam data mining, pengelompokan data juga bisa dilakukan. Tujuannya adalah agar kita dapat mengetahui pola universal data-data yang ada. Anomali data transaksi juga perlu dideteksi untuk dapat mengetahui tindak lanjut berikutnya yang dapat diambil. Semua hal tersebut bertujuan mendukung kegiatan operasional institusi sehingga tujuan akhir institusi diharapkan dapat tercapai.

2.2 Teknik Single Linkage

Single Linkage adalah proses pengklasteran yang didasarkan pada jarak terdekat antar abjeknya (*minimum distance*). Metode *single linkage* sangat bagus untuk melakukan analisa pada tiap tahap pembentukan *cluster*. Metode ini juga sangat cocok untuk dipakai pada kasus *shape independent clustering* karena kemampuannya untuk membentuk *pattern*/pola tertentu dari *cluster*.

Teknik *single linkage* menggunakan prinsip jarak minimum yang diawali dengan mencari dua objek terdekat dan keduanya membentuk *cluster* pertama.

Berikut rumus-rumus yang digunakan untuk menyelesaikan *clustering* menggunakan teknik *single linkage* dengan pendekatan *Euclidean distance* yaitu:

1. Menghitung Nilai Rata-Rata Variabel

$$x = \frac{\sum_{i=1}^n x_i}{n} \dots\dots\dots (1)$$

Dimana:

x = Nilai rata-rata variabel

x_2 = Nilai variabel

Π = jumlah atau kuantitas objek

2. Menghitung Nilai Standar Deviasi

$$S(x) = \sqrt{\frac{\sum_{i=1}^n (x_i - x)^2}{n-1}} \dots\dots\dots (2)$$

Dimana:

$s(x)$ = Nilai standar deviasi

x_2 = Nilai variabel

x = Nilai rata-rata variabel

Π = Jumlah atau kuantitas objek

3. Menghitung Nilai Zero Standar

$$z_i = \frac{x_i - x}{s(x)} \dots\dots\dots (3)$$

Dimana:

Z_i = Nilai Zero Standar

x_i = Nilai variabel

x = Nilai rata-rata variabel

$s(x)$ = Nilai Standar Deviasi

2.3 Mahasiswa

Secara umum, pengertian Mahasiswa merupakan seseorang yang belajar dibangku perkuliahan dengan mengambil jurusan yang disenangi sekaligus jurusan yang di dalamnya ada kemungkinan besar untuk mengembangkan bakatnya. Menurut KBBI (Kamus Besar Bahasa Indonesia), Mahasiswa adalah pelajar perguruan tinggi serta dalam struktur pendidikan Indonesia menduduki jenjang satuan pendidikan tertinggi diantara yang lainnya. Mahasiswa merupakan insan-insan calon sarjana yang terlibat dalam suatu instansi perguruan tinggi, dididik serta diharapkan menjadi calon-calon intelektual [3]-[4].

3. HASIL DAN PEMBAHASAN

Pada proses penentuan pengambilan keputusan untuk pengelompokan bidang skripsi di STMIK Pelita Nusantara Medan, permasalahannya adalah pada penganalisaan data nilai beberapa mata kuliah yang menjadi indikator pembidangan skripsi mahasiswa program studi Teknik Informatika yang telah diperoleh dari *database* STMIK Pelita Nusantara Medan agar dikelompokkan ke dalam sebuah *cluster*, yang mana salah satu tujuannya adalah untuk mengarahkan mahasiswa dalam hal memilih bidang apa yang sesuai untuk dipilih dalam pembuatan skripsi.

Untuk memecahkan permasalahan tersebut, akan dibuat sebuah sistem yang dapat digunakan dalam pengklasteran mahasiswa yang nantinya dapat membantu Ketua Program Studi Teknik Informatika STMIK Pelita Nusantara Medan dalam hal memberikan arahan mengenai pembidangan skripsi kepada mahasiswa. Pemecahan masalah tersebut menggunakan metode *data mining* dengan algoritma teknik *single linkage*.

Adapun yang menjadi sampel pada penelitian ini adalah 20 mahasiswa semester VIII jurusan Teknik Informatika STMIK Pelita Nusantara Medan yang dipilih secara berurut.

Pada metode *hierarchichal clustering*, sebelum dilakukan pengelompokan, setiap data yang ada diasumsikan sebagai *cluster*. Dalam hal ini, jika terdapat jumlah data sebanyak n dan k dianggap sebagai *cluster*, maka besar $n = k$. Kemudian, untuk melakukan pengelompokan dengan tujuan menghasilkan suatu data yang terkelompok adalah dengan cara standarisasi data yang akan dikelompokkan. Hal ini dilakukan agar data mempunyai skala yang samasehingga akan stabil dan kemudian menentukan ukuran kemiripan atau ketidakmiripan antardengan metode pengukuran jarak Euclidean Distance.

Dari hasil perhitungan tersebut, dipilih jarak yang minimal dan digabungkan sehingga besarnya $n = n - 1$. Hal ini terus dilakukan dan akan berhenti sampai memenuhi kondisi jumlah $k = 1$. Pada akhirnya, tahap *hierarchichal clustering* ini akan diperoleh pengelompokan yang diinginkan. Sebelum melakukan perhitungan, data mahasiswa beserta nilai 7 mata kuliah yang diperoleh dari BPN STMIK Pelita Nusantara dapat dilihat pada lampiran 1.

Tabel 1. Data Mahasiswa dan Variabelnya

No	NIM	Nama	x_1	x_2	x_3	x_4	x_5	x_6	x_7	Bidang Study Skripsi
1	130121178	Abu Syahrin Daulay	80,10	81,55	75,75	72,25	81,42	74,00	75,75	SPK
2	130121225	Ade Setiawan	80,10	82,25	72,25	75,00	87,17	74,00	75,75	Sistem Pakar
3	130121218	Adesyah Putra W S	75,00	83,50	84,10	75,50	80,67	75,00	76,90	Data Mining
4	130121081	Agung Sitanggang	76,00	86,50	85,35	75,00	83,70	74,50	77,75	Kriptografi
5	1301211085	Agustina Theresia Butar-butur	79,50	85,75	80,15	71,75	87,33	74,65	82,25	Sistem Pakar
6	130121033	Aldo Paul Gabriel Sitepu	77,10	74,00	80,05	75,50	80,08	76,80	85,65	Kriptografi
7	130121025	Alexander	79,33	81,75	78,50	72,25	80,83	71,83	79,00	SPK
8	130121160	Amagia Br.Ginting	74,00	83,65	73,25	75,25	72,92	70,50	83,00	Sistem Pakar
9	130121219	Andi Lutfy Suhendiko	81,40	78,75	80,75	79,75	75,58	67,62	75,80	Kriptografi
10	130121077	Andreas Gurusinga	79,10	75,33	80,75	78,75	76,58	67,92	79,75	Sistem Pakar
11	130121001	Andri Dirgantara	79,80	89,50	85,80	72,25	87,00	77,75	75,10	Pengolahan Citra
12	130121062	Angelia Mariyani Manurung	75,30	80,70	81,15	73,00	70,67	67,50	75,05	Sistem pakar
13	130121230	Apriadi Syahputra	75,00	86,00	79,75	73,67	90,50	74,67	81,90	Sistem Pakar
14	130121170	Susanti	78,70	82,50	77,80	80,25	95,92	82,55	82,50	SPK
15	130121165	Aprimen Sidauruk	80,90	74,75	79,85	72,25	80,50	71,10	74,75	SPK
16	130121141	Aristina Tafonao	50,65	70,10	77,15	69,50	70,95	72,25	67,00	JST
17	130121152	Aruani Laia	75,75	72,25	77,00	74,25	79,10	71,13	75,25	SPK
18	130121083	Astrid Yolanda Veronika S	81,50	74,00	80,50	75,25	81,25	77,20	75,25	Pengolahan Citra
19	130121126	Aswir Hasibuan	84,00	75,33	73,25	68,75	79,50	73,25	82,70	Kriptografi
20	130121177	Basmen Siregar	84,30	78,75	84,25	76,50	84,50	74,50	84,55	Data Mining

Untuk perhitungan seterusnya, inisial nama mahasiswa dimisalkan dengan huruf A, B, C, D, ..., T dan variabel dimisalkan dengan kode $x_1, x_2, x_3, x_4, x_5, x_6$, dan x_7 , begitu juga dengan nilai asumsi adalah sebagai perumpamaan untuk mempermudah perhitungan.

Tabel 2. Data Sampel dan Variabelnya

No	Objek	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇
1	A	4	4	3	3	4	3	3
2	B	4	4	3	3	4	3	3
3	C	3	4	4	3	4	3	3
4	D	3	4	4	3	4	3	3
5	E	3	4	4	3	4	3	4
6	F	3	3	4	3	4	3	4
7	G	3	4	3	3	4	3	3
8	H	3	4	3	3	3	3	4
9	I	4	3	4	3	3	2	3
10	J	3	3	4	3	3	2	3
11	K	3	4	4	3	4	3	3
12	L	3	4	4	3	3	2	3
13	M	3	4	3	3	5	3	4
14	N	3	4	3	4	5	4	4
15	O	4	3	3	3	4	3	3
16	P	1	3	3	2	3	3	2
17	Q	3	3	3	3	3	3	3
18	R	4	3	4	3	4	3	3
19	S	4	3	3	2	3	3	4
20	T	4	3	4	3	4	3	4

Algoritma Sistem adalah tahapan-tahapan penyelesaian menggunakan data mining dengan teknik single linkage. Adapun langkah-langkah penyelesaiannya adalah sebagai berikut:

1. Melakukan standarisasi data untuk setiap variabel pada masing-masing mahasiswa. Untuk mahasiswa A pada variabel 1 (yang selanjutnya disebut x_1), maka standarisasi dapat dilakukan sebagai berikut:

- a. Mencari nilai rata-rata dari data x_1

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

$$x_1 = \frac{x_{11}+x_{12}+x_{13}+x_{14}+x_{15}+x_{16}+x_{17}+x_{18}+x_{19}+x_{110}+x_{111}+x_{112}+x_{113}+x_{114}+x_{115}+x_{116}+x_{117}+x_{118}+x_{119}+x_{120}}{n}$$

$$x_1 = \frac{4+4+3+3+3+3+3+4+3+3+3+3+4+1+3+4+4+4}{20}$$

$$x_1 = \frac{65}{20}$$

$$x_1 = 3,25$$

- b. Mencari Nilai Standar Deviasi Data x_1

$$S(x) = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}$$

$$\text{Std}(x_1) = \sqrt{\frac{(x_{11}-x_1)^2+(x_{12}-x_1)^2+(x_{13}-x_1)^2+(x_{14}-x_1)^2+(x_{15}-x_1)^2+(x_{16}-x_1)^2+(x_{17}-x_1)^2+(x_{18}-x_1)^2+(x_{19}-x_1)^2+(x_{110}-x_1)^2+(x_{111}-x_1)^2+(x_{112}-x_1)^2+(x_{113}-x_1)^2+(x_{114}-x_1)^2+(x_{115}-x_1)^2+(x_{116}-x_1)^2+(x_{117}-x_1)^2+(x_{118}-x_1)^2+(x_{119}-x_1)^2+(x_{120}-x_1)^2}{n-1}}$$

$$\text{Std}(x_1) = \sqrt{\frac{(4-3,2)^2+(4-3,2)^2+(3-3,2)^2+(3-3,2)^2+(3-3,2)^2+(3-3,2)^2+(3-3,2)^2+(3-3,2)^2+(3-3,2)^2+(4-3,2)^2+(3-3,2)^2+(3-3,2)^2+(3-3,2)^2+(4-3,2)^2+(1-3,2)^2+(3-3,2)^2+(4-3,2)^2+(4-3,2)^2}{20-1}}$$

$$\text{Std}(x_1) = \sqrt{\frac{9,75}{19}}$$

$$\text{Std}(x_1) = 0,716$$

- c. Mencari Data Standar (Nilai Zero Standard) untuk x_1

$$z_i = \frac{x_i - \bar{x}}{s - (x)}$$

$$Z(x_1) = \frac{x_{11} - x_1}{s_i - x_1}$$

$$Z(x_{11}) = \frac{4-3,25}{0,716}$$

$$Z(x_{11}) = 1,047$$

Proses diatas dilakukan berulang hingga sampai dengan x5 baik nilai standarisasi variable, standar deviasi data ataupun nilai zero standar. Setelah semua data distandarkan dengan carayang sama, maka diperoleh tabel baru yang berisi data yang telah standar, yaitu:

Tabel 3. Data Standar

No	OBJEK	x1	x2	x3	x4	x5	x6	x7
1	A	1.047	0.882	-0.975	0.127	0.391	0.224	-0.525
2	B	1.047	0.882	-0.975	0.127	0.391	0.224	-0.525
3	C	-0.349	0.882	0.975	0.127	0.391	0.224	-0.525
4	D	-0.349	0.882	0.975	0.127	0.391	0.224	-0.525
5	E	-0.349	0.882	0.975	0.127	0.391	0.224	1.225
6	F	-0.349	-1.078	0.975	0.127	0.391	0.224	1.225
7	G	-0.349	0.882	-0.975	0.127	0.391	0.224	-0.525
8	H	-0.349	0.882	-0.975	0.127	-1.174	0.224	1.225
9	I	1.047	-1.078	0.975	0.127	-1.174	-2.012	-0.525
10	J	-0.349	-1.078	0.975	0.127	-1.174	-2.012	-0.525
11	K	-0.349	0.882	0.975	0.127	0.391	0.224	-0.525
12	L	-0.349	0.882	0.975	0.127	-1.174	-2.012	-0.525
13	M	-0.349	0.882	-0.975	0.127	1.957	0.224	1.225
14	N	-0.349	0.882	-0.975	2.665	1.957	2.460	1.225
15	O	1.047	-1.078	-0.975	0.127	0.391	0.224	-0.525
16	P	-3.141	-1.078	-0.975	-2.411	-1.174	0.224	-2.276
17	Q	-0.349	-1.078	-0.975	0.127	-1.174	0.224	-0.525
18	R	1.047	-1.078	0.975	0.127	0.391	0.224	-0.525
19	S	1.047	-1.078	-0.975	-2.411	-1.174	0.224	1.225
20	T	1.047	-1.078	0.975	0.127	0.391	0.224	1.225

- d. Membuat matriks jarak antar objek dengan menggunakan rumus jarak *Euclidean Distance*. Caranya adalah dengan menggunakan rumus :

$$d(x,y) = \sqrt{\sum (xi - yi)^2} = \sqrt{\sum (x1 - y1)^2 + (x2 - y2)^2 + \dots + (xn - yn)^2}$$

Misal, untuk mencari jarak kedekatan antara mahasiswa A dan B

$$d(A,B) = \sqrt{(A1 - B1)^2 + (A2 - B2)^2 + (A3 - B3)^2 + (A4 - B4)^2 + (A5 - B5)^2 + (A6 - B6)^2 + (A7 - B7)^2}$$

$$d(A,B) = \sqrt{1.047 - 1.047)^2 + (0.755 - 0.755)^2 + (-0.872 - (-0.872))^2 + (0.100 - 0.100)^2 + (0.305 - 0.305)^2 + (0.174 - 0.174)^2 + (-0.523 - (-0.523))^2}$$

$$d(A,B) = \sqrt{0}$$

$$d(A,B) = 0$$

3.1 Penerapan Teknik Single Linkage Clustering

- Melakukan Pengelompokan dengan Euclidean Single Linkage Clustering

- Mencari Nilai Terkecil dari Matriks Jarak

Objek A, B, C, D, dan K mempunyai nilai terkecil yaitu 0,000. Maka objekA, B, C, D, dan K bergabung menjadi satu cluster yaitu cluster 1.

- Menghitung Jarak Antar Cluster ABCDK dengan objek lainnya di luar cluster 1 yaitu: E, F, G, H, I, J, L, M, N, O, P, Q, R, S, T.

- $d(ABCDK) E = \min \{dAE, dBE, dCE, dDE, dKE\} = 1,751$
- $d(ABCDK) F = \min \{dAF, dBF, dCF, dDF, dKF\} = 2,627$
- $d(ABCDK) G = \min \{dAG, dBG, dCG, dDG, dKG\} = 1,396$
- $d(ABCDK) H = \min \{dAH, dBH, dCH, dDH, dKH\} = 2,732$
- $d(ABCDK) I = \min \{dAI, dBI, dCI, dDI, dKI\} = 3,639$
- $d(ABCDK) J = \min \{dAJ, dBJ, dCJ, dDJ, dKJ\} = 3,360$
- $d(ABCDK) L = \min \{dAL, dBL, dCL, dDL, dKL\} = 2,730$
- $d(ABCDK) M = \min \{dAM, dBM, dCM, dDM, dKM\} = 2,732$
- $d(ABCDK) N = \min \{dAN, dBN, dCN, dDN, dKN\} = 4,348$
- $d(ABCDK) O = \min \{dAO, dBO, dCO, dDO, dKO\} = 1,959$
- $d(ABCDK) P = \min \{dAP, dBP, dCP, dDP, dKP\} = 5,234$
- $d(ABCDK) Q = \min \{dAQ, dBQ, dCQ, dDQ, dKQ\} = 2,870$
- $d(ABCDK) R = \min \{dAR, dBR, dCR, dDR, dKR\} = 2,406$
- $d(ABCDK) S = \min \{dAS, dBS, dCS, dDS, dKS\} = 3,974$
- $d(ABCDK) T = \min \{dAT, dBT, dCT, dDT, dKT\} = 2,975$

- c. Mencari Nilai Terkecil dari Matriks Jarak Cluster 1
Objek ABCDK, F, G, I, J dan T mempunyai nilai terkecil yaitu 1,396. Maka objek ABCDK, F, G, I, J dan T bergabung menjadi satu cluster yaitu cluster 2.
- d. Menghitung Jarak Antar Cluster ABCDK-FGIJT dengan objek lainnya di luar cluster 2 yaitu: E, H, L, M, N, O, P, Q, R, S.
 1. $d(\text{ABCDK-FGIJT}) E = \min \{d(\text{ABCDK-E}, d_{FE}, d_{GE}, d_{IE}, d_{JE}, d_{TE})\} = 1,751$
 2. $d(\text{ABCDK-FGIJT}) H = \min \{d(\text{ABCDK-H}, d_{FH}, d_{GH}, d_{IH}, d_{JH}, d_{TH})\} = 2,349$
 3. $d(\text{ABCDK-FGIJT}) L = \min \{d(\text{ABCDK-L}, d_{FL}, d_{GL}, d_{IL}, d_{JL}, d_{TL})\} = 1,959$
 4. $d(\text{ABCDK-FGIJT}) M = \min \{d(\text{ABCDK-M}, d_{FM}, d_{GM}, d_{IM}, d_{JM}, d_{TM})\} = 2,349$
 5. $d(\text{ABCDK-FGIJT}) N = \min \{d(\text{ABCDK-N}, d_{FN}, d_{GN}, d_{IN}, d_{JN}, d_{TN})\} = 4,118$
 6. $d(\text{ABCDK-FGIJT}) O = \min \{d(\text{ABCDK-O}, d_{FO}, d_{GO}, d_{IO}, d_{JO}, d_{TO})\} = 1,959$
 7. $d(\text{ABCDK-FGIJT}) P = \min \{d(\text{ABCDK-P}, d_{FP}, d_{GP}, d_{IP}, d_{JP}, d_{TP})\} = 4,857$
 8. $d(\text{ABCDK-FGIJT}) Q = \min \{d(\text{ABCDK-Q}, d_{FQ}, d_{GQ}, d_{IQ}, d_{JQ}, d_{TQ})\} = 2,508$
 9. $d(\text{ABCDK-FGIJT}) R = \min \{d(\text{ABCDK-R}, d_{FR}, d_{GR}, d_{IR}, d_{JR}, d_{TR})\} = 1,751$
 10. $d(\text{ABCDK-FGIJT}) S = \min \{d(\text{ABCDK-S}, d_{FS}, d_{GS}, d_{IS}, d_{JS}, d_{TS})\} = 3,563$
- e. Mencari Nilai Terkecil dari Matriks Jarak Cluster 2
Objek ABCDKFGIJT, E dan R mempunyai nilai terkecil yaitu 1,751. Maka objek ABCDKFGIJT, E dan R bergabung menjadi satu cluster yaitu cluster 3.
- f. Menghitung jarak antar cluster ABCDKFGIJT-ER dengan objek lainnya di luar cluster 3 yaitu: H, L, M, N, O, P, Q, S.
 1. $d(\text{ABCDKFGIJT-ER}) H = \{d(\text{ABCDKFGIJT-H}, d_{EH}, d_{RH})\} = 2,349$
 2. $d(\text{ABCDKFGIJT-ER}) L = \{d(\text{ABCDKFGIJT-L}, d_{EL}, d_{RL})\} = 1,959$
 3. $d(\text{ABCDKFGIJT-ER}) M = \{d(\text{ABCDKFGIJT-M}, d_{EM}, d_{RM})\} = 2,349$
 4. $d(\text{ABCDKFGIJT-ER}) N = \{d(\text{ABCDKFGIJT-N}, d_{EN}, d_{RN})\} = 4,118$
 5. $d(\text{ABCDKFGIJT-ER}) O = \{d(\text{ABCDKFGIJT-O}, d_{EO}, d_{RO})\} = 1,949$
 6. $d(\text{ABCDKFGIJT-ER}) P = \{d(\text{ABCDKFGIJT-P}, d_{EP}, d_{RP})\} = 4,857$
 7. $d(\text{ABCDKFGIJT-ER}) Q = \{d(\text{ABCDKFGIJT-Q}, d_{EQ}, d_{RQ})\} = 2,508$
 8. $d(\text{ABCDKFGIJT-ER}) S = \{d(\text{ABCDKFGIJT-S}, d_{ES}, d_{RS})\} = 3,563$
- g. Mencari Nilai Terkecil dari Matriks Jarak Cluster 3
Objek ABCDKFGIJTER dan O mempunyai nilai terkecil yaitu 1,949. Maka objek ABCDKFGIJTER dan O bergabung menjadi satu cluster yaitu cluster 4.
- h. Menghitung jarak antar cluster ABCDKFGIJTER-O dengan objek lainnya di luar cluster 4 yaitu: H, L, M, N, P, Q, S.
 1. $d(\text{ABCDKFGIJTER-O}) H = \{d(\text{ABCDKFGIJTER-H}, d_{OH})\} = 2,349$
 2. $d(\text{ABCDKFGIJTER-O}) L = \{d(\text{ABCDKFGIJTER-L}, d_{OL})\} = 1,959$
 3. $d(\text{ABCDKFGIJTER-O}) M = \{d(\text{ABCDKFGIJTER-M}, d_{OM})\} = 2,349$
 4. $d(\text{ABCDKFGIJTER-O}) N = \{d(\text{ABCDKFGIJTER-N}, d_{ON})\} = 4,118$
 5. $d(\text{ABCDKFGIJTER-O}) P = \{d(\text{ABCDKFGIJTER-P}, d_{OP})\} = 4,857$
 6. $d(\text{ABCDKFGIJTER-O}) Q = \{d(\text{ABCDKFGIJTER-Q}, d_{OQ})\} = 2,098$
 7. $d(\text{ABCDKFGIJTER-O}) S = \{d(\text{ABCDKFGIJTER-S}, d_{OS})\} = 3,458$
- i. Mencari Nilai Terkecil dari Matriks Jarak Cluster 4
Objek ABCDKFGIJTERO dan L mempunyai nilai terkecil yaitu 1,959. Maka objek ABCDKFGIJTERO dan L bergabung menjadi satu cluster yaitu cluster 5.

- j. Menghitung jarak antar *cluster* ABCDKFGIJTERO-L dengan objek lainnya di luar *cluster* 5 yaitu: H, M, N, P, Q, S.
1. $d(\text{ABCDKFGIJTERO-L}) H = \{d(\text{ABCDKFGIJTERO-H}, d_{LH}) = 2,349$
 2. $d(\text{ABCDKFGIJTERO-L}) M = \{d(\text{ABCDKFGIJTERO-M}, d_{LM}) = 2,349$
 3. $d(\text{ABCDKFGIJTERO-L}) N = \{d(\text{ABCDKFGIJTERO-N}, d_{LN}) = 4,118$
 4. $d(\text{ABCDKFGIJTERO-L}) P = \{d(\text{ABCDKFGIJTERO-P}, d_{LP}) = 4,857$
 5. $d(\text{ABCDKFGIJTERO-L}) Q = \{d(\text{ABCDKFGIJTERO-Q}, d_{LQ}) = 2,098$
 6. $d(\text{ABCDKFGIJTERO-L}) S = \{d(\text{ABCDKFGIJTERO-S}, d_{LS}) = 3,458$
- k. Mencari Nilai Terkecil dari Matriks Jarak Cluster 5
Objek ABCDKFGIJTEROL dan Q mempunyai nilai terkecil yaitu 2,098. Maka objek ABCDKFGIJTEROL dan Q bergabung menjadi satu cluster yaitu cluster 6.
- l. Menghitung jarak antar cluster ABCDKFGIJTEROL-Q dengan objek lainnya di luar cluster 6 yaitu: H, M, N, P, S.
1. $d(\text{ABCDKFGIJTEROL-Q}) H = \{d(\text{ABCDKFGIJTEROL-H}, d_{QH}) = 2,349$
 2. $d(\text{ABCDKFGIJTEROL-Q}) M = \{d(\text{ABCDKFGIJTEROL-M}, d_{QM}) = 2,349$
 3. $d(\text{ABCDKFGIJTEROL-Q}) N = \{d(\text{ABCDKFGIJTEROL-N}, d_{QN}) = 4,118$
 4. $d(\text{ABCDKFGIJTEROL-Q}) P = \{d(\text{ABCDKFGIJTEROL-P}, d_{QP}) = 4,159$
 5. $d(\text{ABCDKFGIJTEROL-Q}) S = \{d(\text{ABCDKFGIJTEROL-S}, d_{QS}) = 3,384$
- m. Mencari Nilai Terkecil dari Matriks Jarak Cluster 6
Objek ABCDKFGIJTEROLQ, H dan M mempunyai nilai terkecil yaitu 2,349. Maka objek ABCDKFGIJTEROLQ, H dan M bergabung menjadi satu cluster yaitu cluster 7.
- n. Menghitung jarak antar *cluster* ABCDKFGIJTEROLQ-HM dengan objek lainnya di luar *cluster* 7 yaitu: N, P, S
1. $d(\text{ABCDKFGIJTEROLQ-HM}) N = \{d(\text{ABCDKFGIJTEROLQ-N}, d_{HN}, d_{MN}) = 3,382$
 2. $d(\text{ABCDKFGIJTEROLQ-HM}) P = \{d(\text{ABCDKFGIJTEROLQ-P}, d_{HP}, d_{MP}) = 4,159$
 3. $d(\text{ABCDKFGIJTEROLQ-HM}) S = \{d(\text{ABCDKFGIJTEROLQ-S}, d_{HS}, d_{MS}) = 3,458$
- o. Mencari Nilai Terkecil dari Matriks Jarak Cluster 7
Objek ABCDKFGIJTEROLQHM dan N mempunyai nilai terkecil yaitu 3,382. Maka objek ABCDKFGIJTEROLQHM dan N bergabung menjadi satu cluster yaitu cluster 8.
- p. Menghitung jarak antar *cluster* ABCDKFGIJTEROLQHM-N dengan objeklainnya di luar *cluster* 8 yaitu: P, S
1. $d(\text{ABCDKFGIJTEROLQHM-N}) P = \{d(\text{ABCDKFGIJTEROLQHM-P}, d_{NP}) = 4,159$
 2. $d(\text{ABCDKFGIJTEROLQHM-N}) S = \{d(\text{ABCDKFGIJTEROLQHM-S}, d_{NS}) = 3,458$
- q. Mencari Nilai Terkecil dari Matriks Jarak Cluster 8
Objek ABCDKFGIJTEROLQHMN dan S mempunyai nilai terkecil yaitu 3,458. Maka objek ABCDKFGIJTEROLQHMN dan S bergabung menjadi satu cluster yaitu cluster 9.
- r. Menghitung jarak antar *cluster* ABCDKFGIJTEROLQHMN-S dengan objek lainnya di luar *cluster* 9 yaitu: P
 $d(\text{ABCDKFGIJTEROLQHMN-S})P = \min\{d(\text{ABCDKFGIJTEROLQHMN-P}, d_{SP}) = 4,159$
- Setelah proses pembentukan cluster dilakukan, berikut adalah data keseluruhan dari cluster-cluster yang sudah terbentuk:

Tabel 4. Data Keseluruhan Cluster

Cluster	Kode Cluster	Keterangan	Distance
1	ABCDK	Abu Syahrin Daulay(130121178), Ade Setiawan (130121225), Adesyah PutraW S (130121218), Agung Sitanggang (130121081), Andri Dirgantara (130121001).	0,000
2	ABCDK-FGIJT	Abu Syahrin Daulay(130121178), Ade Setiawan (130121225), Adesyah PutraW S (130121218), Agung Sitanggang (130121081), Andri Dirgantara (130121001), Aldo Paul Gabriel Sitepu (130121033), Alexander (130121025), Andi Lutfy Suhendiko (130121219), Andreas Gurusinga (130121077), Basmen Siregar (130121177).	1,396
3	ABCDKFGIJT-ER	Abu Syahrin Daulay(130121178), Ade Setiawan (130121225), Adesyah PutraW S (130121218), Agung Sitanggang (130121081), Andri Dirgantara	

Cluster	Kode Cluster	Keterangan	Distance
8	ABCDKFGIJTEROL QHM-N	Abu Syahrin Daulay(130121178), Ade Setiawan (130121225), Adesyah Putra W S (130121218), Agung Sitanggang (130121081), Andri Dirgantara (130121001), Aldo Paul Gabriel Sitepu (130121033), Alexander (130121025), Andi Lutfy Suhendiko (130121219), Andreas Gurusinga (130121077), Basmen Siregar (130121177), Agustina Theresia Butar-butur (130121085), Astrid Yolanda Veronika S (130121083), Aprimen Sidaauruk (130121165), Angelia Mariyani Manurung (130121062), Aruani Laila (130121152), Amagia Br.Ginting (130121160), Apriadi Syahputra (130121230), Aprida Susanti (130121230).	3,382
9	ABCDKFGIJTEROL QHMNS	Abu Syahrin Daulay(130121178), Ade Setiawan (130121225), Adesyah Putra W S (130121218), Agung Sitanggang (130121081), Andri Dirgantara (130121001), Aldo Paul Gabriel Sitepu (130121033), Alexander (130121025), Andi Lutfy Suhendiko (130121219), Andreas Gurusinga (130121077), Basmen Siregar (130121177), Agustina Theresia Butar-butur (130121085), Astrid Yolanda Veronika S (130121083), Aprimen Sidaauruk (130121165), Angelia Mariyani Manurung (130121062), Aruani Laila (130121152), Amagia Br.Ginting (130121160), Apriadi Syahputra (130121230), Aprida Susanti (130121230), Aswir Hasibuan (130121126).	3.458

Adapun mahasiswa yang dibagi berdasarkan hasil cluster-cluster di atas adalah sebagai berikut:

Tabel 5. Hasil Cluster

Nomor Cluster	NIM	Nama	Judul Skripsi	Ya / Tidak
0	130121141	Aristina Tafonao	JST	Ya
1	130121178	Abu Syahrin Daulay	SPK	Tidak
	130121225	Ade Setiawan	Sistem Pakar	Tidak
	130121218	Adesyah Putra W S	Data Mining	Tidak
	130121081	Agung Sitanggang	Kriptografi	Ya
	130121001	Andri Digantara	Pengolahan Citra	Ya
	130121033	Aldo Paul Gabriel Sitepu	Kriptografi	Ya
	130121025	Alexander	SPK	Tidak
	130121219	Andi LutfySuhendiko	Kriptografi	Ya
	130121077	Andreas Gurusinga	Sistem Pakar	Tidak
	130121177	Basmen Siregar	Data Mining	Tidak
	130121085	Agustina Theresia Butar-butur	Sistem Pakar	Tidak
	130121083	Astrid Yolanda Veronika S	Pengolahan Citra	Ya
	130121165	Aprimen Sidaauruk	SPK	Tidak
	130121062	Angelia Mariyani Manurung	Sistem Pakar	Tidak
	130121152	Aruani laila	SPK	Tidak
	130121160	Amagia Br.Ginting	Sistem Pakar	Tidak

Nomor Cluster	NIM	Nama	Judul Skripsi	Ya / Tidak
	130121230	Apriadi Syahputra	Sistem Pakar	Tidak
	130121170	Aprida Susanti	SPK	Tidak
	130121126	Aswir Hasibuan	Kriptografi	Ya

Berdasarkan data 7 mata kuliah yang menjadi indikator pembidangan skripsi pada Tabel 5. maka dapat disimpulkan bahwa mahasiswa yang beradapada cluster 0 adalah mahasiswa yang diarahkan untuk mengambil bidang skripsi Sistem pendukung Keputusan, Sistem Pakar, Data Mining dan Jaringan Syaraf Tiruan karena nilai mata kuliah yang termasuk di bidang tersebut cenderung lebih tinggi dibandingkan dengan nilai mata kuliah lain sedangkan mahasiswa yang berada di cluster 1 adalah mahasiswa yang diarahkan untuk mengambil bidang skripsi: Pemodelan dan Simulasi, Pengolahan Citra dan Keamanan Komputer karena nilai mata kuliah tersebut cenderung lebih bagus dibandingkan dengan nilai mata kuliah pada cluster 0.

4. KESIMPULAN

Dari hasil penelitian disimpulkan bahwa teknik *single linkage* dapat digunakan untuk menyelesaikan permasalahan yang berkenaan dengan pembidangan skripsi mahasiswa di STMIK Pelita Nusantara Medan. Penerapan teknik *single linkage* di STMIK Pelita Nusantara Medan dapat membantu Ketua Program Studi khususnya Program Teknik Informatika dalam memberikan arahan kepada mahasiswa dalam penentuan bidang ilmu yang akan dipilih dalam penyusunan judul skripsi.

REFERENCES

- [1] R. Handoyo et al., "Perbandingan Metode Clustering Menggunakan Metode Single Linkage Dan K-Means Pada Pengelompokan Dokumen," JSM STMIK Mikroskil, vol. 15, no. 2, pp. 73–82, 2014.
- [2] F. N. Adhiatma et al., "PERANCANGAN DAN ANALISIS CLUSTERING DATA MENGGUNAKAN METODE SINGLE LINKAGE UNTUK BERITA BERBAHASA INGGRIS DESIGN AND ANALYSIS OF DATA CLUSTERING USING SINGLE LINKAGE," vol. 3, no. 2, pp. 2285–2291, 2016.
- [3] J. Ilmiah, I. Komputa, E. Volume, B. Issn, and J. D. Bandung, "PENERAPAN DATA MINING UNTUK MEMBENTUK KELOMPOK BELAJAR MENGGUNAKAN METODE CLUSTERING DI SMPN 19 BANDUNG Andre Catur Prasetyo Jurnal Ilmiah Komputer dan Informatika (KOMPUTA)."
- [4] S. Kasus, U. Dehasen, S. Haryati, A. Sudarsono, and E. Suryana, "IMPLEMENTASI DATA MINING UNTUK MEMPREDIKSI MASA STUDI MAHASISWA MENGGUNAKAN ALGORITMA C4.5," vol. 11, no. 2, pp. 130–138, 2015.
- [5] S. Haryati, A. Sudarsono, and E. Suryana, "Implementasi Data Mining Untuk Memprediksi Masa Studi Mahasiswa Menggunakan Algoritma C4.5 (Studi Kasus: Universitas Dehasen Bengkulu)," J. Media Infotama Vol., vol. 11, no. 2, pp. 130–138, 2015.
- [6] R. Handoyo et al., "Perbandingan Metode Clustering Menggunakan Metode Single Linkage Dan K-Means Pada Pengelompokan Dokumen," JSM STMIK Mikroskil, vol. 15, no. 2, pp. 73–82, 2014.
- [7] J. Hartono, "Analisis Dan Desain Sistem Informasi," II., Yogyakarta: Andi, 1999, p. 795.
- [8] D. P. Utomo and B. Purba, "Penerapan Datamining pada Data Gempa Bumi Terhadap Potensi Tsunami di Indonesia," Prosiding Seminar Nasional Riset Information Science (SENARIS), vol. 1, pp. 846–853, 2019.
- [9] R. Amelia and D. P. Utomo, "ANALISA POLA PEMESANAN PRODUK MODERN TRADE INDEPENDENT DENGAN MENEREPAKAN ALGORITMA FP. GROWTH (STUDI KASUS: PT. ADAM DANI LESTARI)," KOMIK (Konferensi Nasional Teknologi Informasi dan Komputer), vol. 3, no. 1, pp. 416–423, 2019.
- [10] D. P. Utomo and M. Mesran, "Analisis Komparasi Metode Klasifikasi Data Mining dan Reduksi Atribut Pada Data Set Penyakit Jantung," JURNAL MEDIA INFORMATIKA BUDIDARMA, vol. 4, no. 2, pp. 437–444, 2020.
- [11] M. A. Puspa, "Implementasi Data Mining Klasifikasi Algoritma C4.5 Dalam Perekrutan Perangkat Desa," Building of Informatics, Technology and Science (BITS), vol. 1, no. 2, pp. 92–97, 2019.
- [12] Mariani, "Implementasi Algoritma Apriori Pada Persediaan Barang Minuman Larutan Cap Kaki Tiga (Studi Kasus: PT Adyajati Lestari)," Building Informatics, Technology and Science (BITS), vol. 1, no. 1, pp. 7–27, 2019.