

OPTIMALISASI TEKNIK *SINGLE LINKAGE* DALAM MENGELOMPOKKAN BIBIT SAWIT MENERAPKAN METODE *CLUSTERING*

Ramen Antonov Purba

Manajemen Informatika, Politeknik Unggul LP3M

Jl. Iskandar Muda No. 3 CDEF, Medan, 20153

Telp : (061) 4156355

E-mail : ramenantonovpurba@gmail.com

Abstrak

Dalam budidaya kelapa sawit, bibit faktor pertama yang harus disiapkan. Petani dan pelaku usaha perkebunan kelapa sawit mengalami kesulitan menentukan bibit berkualitas. Kondisi demikian harus dicari solusinya. Sistem berbasis komputer dengan menggunakan data mining dapat menjadi solusi. Data Mining adalah penemuan informasi baru dengan mencari pola atau aturan tertentu dari sejumlah data dalam jumlah besar. Mengingat jumlah bibit sawit cukup banyak dilakukan dengan mengkombinasikan data mining secara clustering. Teknik yang digunakan Single Linkage. Dalam paper ini penulis melakukan pengelompokan bibit sawit. Yang dijadikan sebagai penentu kualitas adalah Baik, Cukup dan Kurang. Kriteria penilaian antara lain : Tinggi bibit, karakteristik daun, dan diameter batang. Diberi range nilai antar 0-10. Hitung passing grade. Mencari Nilai Rata-Rata Tiap Parameter. Mencari Nilai Standar Deviasi Tiap Parameter. Mencari Nilai Standar Untuk Masing-Masing Variabel. Setelah mencapai nilai normalisasi, maka dilakukan perhitungan menggunakan Euclidean Distance. Untuk memudahkan pola kerja akan dibangun aplikasi menggunakan aplikasi berbasis desktop.

Kata kunci : bibit, clustering, data mining, single linkage, sawit, unggul

1. PENDAHULUAN

Luas total kebun kelapa sawit di Indonesia 7,3 juta hektar. Dikelola oleh 12 juta petani sawit. Ada 16 provinsi penghasil kelapa sawit terbesar di Indonesia mulai dari Aceh sampai Papua. Dalam budidaya tanaman kelapa sawit, bibit merupakan faktor pertama yang harus disiapkan. Kualitas bibit menentukan masa depan usaha hingga 30 tahun ke depan [1][2].

Petani dan pelaku usaha perkebunan kelapa sawit seringkali mengalami kesulitan dalam menentukan bibit yang berkualitas. Pernah ada petani yang membuka lahan perkebunan dengan menanam 2500 bibit. Setelah 3 tahun ditanam belum berbuah. Petani terpaksa mengganti seluruh tanaman sawit tersebut [1][2]. Kondisi demikian tentu harus dicari solusinya. Sistem berbasis komputer diyakini mampu memberikan solusi atas permasalahan tersebut.

Dalam rangka mencapai potensi produksi tanaman, pengelolaan bibit sawit dianjurkan untuk melaksanakan sistem pengelompokan penanaman di lapangan. Pengelompokan ini dimaksudkan untuk menjamin keseragaman pertumbuhan tanaman. Konteks kerja demikian dapat menggunakan teknologi berbasis komputerisasi dengan *data mining*. *Data Mining* adalah penambangan atau penemuan informasi baru dengan mencari pola atau aturan tertentu dari sejumlah data dalam jumlah besar [3][4][5]. Mengingat jumlah bibit sawit yang cukup banyak sehingga menimbulkan variasi individu yang cukup banyak, maka dapat dilakukan dengan mengkombinasikan *data mining* secara *clustering*. *Clustering data mining* yang digunakan dalam penelitian adalah *Teknik Single Linkage*. Dimana data-data yang di *cluster* terbentuk berdasarkan masing-masing aturan. Untuk meningkatkan akurasi maka dilakukan *cluster analysis* [6][7][8].

Penelitian yang sebelumnya yang dilakukan PM Booma dan S Prabhakaran dalam penelitiannya berjudul *Classification Of Genes For Disease Identification Using Data Mining Techniques*, melakukan identifikasi

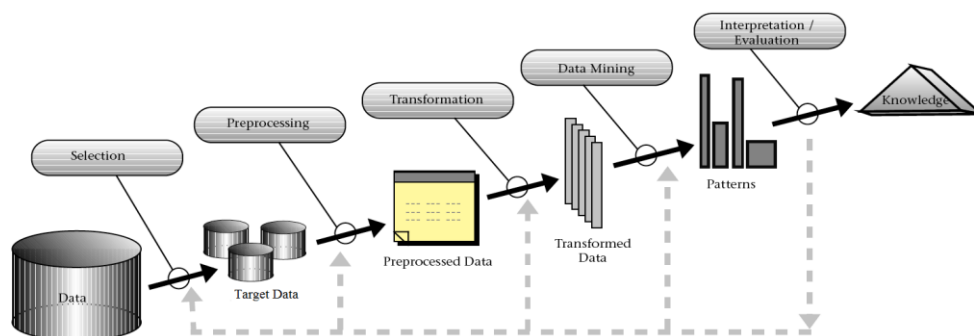
menggunakan teknik *data mining* dalam mengklasifikasi unsur gen biologis. Mengekstraksi informasi proses biologis dari *dataset* ekspresi gen [4]. Jyoti Sinha dan Deepak Shrivastava dalam penelitiannya berjudul *Clustering Based Approach in Data Mining: A Review*, melakukan identifikasi *data mining clustering* terkait metode pembelajaran. Melakukan klasifikasi dalam pembelajaran yang diawasi dan tidak diawasi [5]. S.R.Pande, S.S.Sambare, V.M.Thakre dalam penelitiannya berjudul *Data Clustering Using Data Mining Techniques*, melakukan pengelompokan lokasi rawan gempa bumi [6]. Sankar Rajagopal dalam penelitiannya berjudul *Customer Data Clustering Using Data Mining Technique*, melakukan pengklasifikasian data pelanggan terkait profit yang dapat dihasilkan [7]. Penelitian sebelumnya sama sekali belum ada yang secara spesifik membahas tentang pengelompokan bibit kelapa sawit dalam rangka membantu petani dan pelaku usaha perkebunan kelapa sawit.

Dalam penelitian ini penulis akan melakukan pengelompokan bibit sawit. Hal yang dijadikan sebagai penentu kualitas bibit sawit adalah baik, cukup dan kurang. Kriteria penilaian antara lain tinggi bibit, karakteristik daun, dan diameter batang. Setiap kriteria penilaian diberi *range* nilai antar 0-10. Lalu hitunglah *passing grade* dengan rumus *passing grade*, mencari nilai rata-rata tiap parameter, mencari nilai standar deviasi tiap parameter, dan mencari nilai standar untuk masing-masing variabel. Setelah mendapat nilai normalisasi, maka dilakukan perhitungan menggunakan *Euclidean Distance*.

2. TINJAUAN PUSTAKA

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 [3]. Secara umum *Data Mining* memiliki beberapa kajian. *Data Mining* merupakan pusat dari beberapa kajian, diantaranya adalah estimasi, seleksi variabel, *clustering*, visualisasi, *market basket analysis* dan klasifikasi. Semua kajian tersebut termasuk ke dalam *data mining*. Hubungan yang dicari dalam *Data Mining* dapat berupa hubungan antara dua atau lebih objek dalam satu dimensi yang sama. Misalnya dalam dimensi produk dapat melihat keterkaitan pembelian suatu produk dengan produk yang lain. Selain itu, hubungan juga dapat dilihat antara dua atau lebih atribut dan dua atau lebih objek [8][9][10]. Tahapan *Data Mining* dapat dilihat pada Gambar 1.



Gambar 1. Tahapan Data Mining [11]

2.2 Clustering

Clustering (pengelompokan data) mempertimbangkan sebuah pendekatan penting untuk mencari kesamaan dalam data dan menempatkan data yang sama ke dalam kelompok-kelompok. Gagasan mengenai pengelompokan data, atau *clustering*, memiliki sifat yang sederhana dan dekat dengan cara berpikir manusia kapanpun kepada kita dipresentasikan jumlah data yang besar, kita biasanya cenderung merangkumkan jumlah data yang besar ini ke dalam sejumlah kecil kelompok-kelompok atau kategori-kategori untuk memfasilitasi analisisnya lebih lanjut [10][11][12].

2.3 Teknik Single Linkage

Single Linkage disebut juga dengan *minimum link*, dimana similiaritas dari dua *cluster* didasarkan terhadap dua titik dari dua *cluster* yang berbeda. Kelebihannya dapat menangani bentuk sekelompok *cluster* yang

tidak elips, sedangkan kekurangannya adalah sensitif terhadap *noise* ataupun *outliers*. Adapun algoritma pada *Single Linkage clustering* adalah sebagai berikut [13]:

1. Mulai dengan N *cluster*, setiap *cluster* mengandung entiti tunggal dan sebuah matriks simetrik dari jarak (*similarities*).
2. Cari matriks jarak untuk pasangan *cluster* yang terdekat (paling mirip). Misalkan jarak antara *cluster* X dan Y yang paling mirip.
3. Gabungkan *cluster* X dan Y. Label *cluster* yang baru dibentuk dengan (XY).
4. Ulangi langkah 2 dan 3 sebanyak $(N-1)$ kali hingga semua objek akan berada dalam *cluster* tunggal.

3. PEMBAHASAN

3.1 Analisis Permasalahan

Permasalahan-permasalahan itu antara lain, yaitu :

- a) Sulit melakukan analisa sehingga belum dapat memberikan *output* tentang berapa jenis dan ciri fisik yang dimiliki bibit sawit dan kelompok-kelompok bibit sawit.
- b) Produsen bibit sawit terlalu banyak, sehingga diperlukan sistem yang bisa mendeteksi produsen.

3.2 Algoritma Sistem

Dalam pengelompokan bibit sawit dijadikan penentu kualitas bibit sawit adalah Baik, Cukup dan Kurang. Kriteria penilaian adalah : 1) Tinggi bibit; 2) Karakteristik daun; 3) Diameter batang. Setiap kriteria penilaian diberi *range* nilai antar 0-10.

3.3 Pengujian

Tabel 1 dan Tabel 2 merupakan data hasil observasi bibit sawit yang diambil secara acak dari pembibitan dengan usia bibit sekitar 6 bulan, dimana nantinya data tersebut akan diolah menggunakan metode *Data Mining* dengan teknik *Single Linkage*.

Tabel 1. Data hasil Observasi

No	Bibit	Kriteria	Keterangan	Asumsi
1	A	Tinggi bibit	27 cm	3
		Karakteristik daun	Warna hijau muda, tekstur daun keras dan tidak ada terdapat lubang.	2
		Diameter batang	1 mm	2
2	B	Tinggi bibit	28 cm	2
		Karakteristik daun	Warna hijau muda, tekstur daun keras dan tidak ada terdapat lubang	3
		Diameter batang	2 mm	3
3	C	Tinggi bibit	23 cm	2
		Karakteristik daun	Warna hijau pekat, tekstur daun lentur dan tidak terdapat lubang.	4
		Diameter batang	2 mm	3
4	D	Tinggi bibit	21 cm	1
		Karakteristik daun	Warna kekuningan, tekstur daun lentur tetapi terdapat sedikit lubang.	2
		Diameter batang	0,5 mm	1
5	E	Tinggi bibit	30 cm	4
		Karakteristik daun	Warna kekuningan, tekstur daun lentur tetapi terdapat sedikit lubang.	2
		Diameter batang	0,5 mm	1
6	F	Tinggi bibit	27 cm	3
		Karakteristik daun	Warna kekuningan, tekstur daun lentur tetapi terdapat sedikit lubang.	2
		Diameter batang	1 mm	2
7	G	Tinggi bibit	24 cm	2
		Karakteristik daun	Warna hijau muda, tekstur daun keras dan tidak ada terdapat lubang.	3
		Diameter batang	1 mm	2

8	H	Tinggi bibit	23 cm	2
		Karakteristik daun	Warna hijau pekat, tekstur daun lentur dan tidak terdapat lubang.	4
		Diameter batang	2 mm	3
9	I	Tinggi bibit	20 cm	1
		Karakteristik daun	Warna kekuningan, tekstur daun lentur tetapi terdapat sedikit lubang.	2
		Diameter batang	1 mm	2
10	J	Tinggi bibit	28 cm	3
		Karakteristik daun	Warna kekuningan, tekstur daun keras dan terdapat banyak lubang.	1
		Diameter batang	1 mm	2

Tabel 2. Data Bibit sawit dan Variabelnya

No	Bibit	V1	V2	V3
1	A	3	2	2
2	B	2	3	3
3	C	2	4	3
4	D	1	2	1
5	E	4	2	1
6	F	3	2	2
7	G	2	3	2
8	H	2	4	3
9	I	1	2	2
10	J	3	1	2

Keterangan :

V1 = Tinggi Bibit

V2 = Karakteristik Daun

V3 = Diameter Batang

Dalam pengerjaan menggunakan metode ini, ada beberapa perhitungan yang harus dilakukan diantaranya :

1. Mencari Nilai Rata-Rata Tiap Parameter

$$\bar{X} = \frac{\sum_i^n xi}{n} \quad (1)$$

Dimana X adalah nilai rata-rata parameter, xi adalah nilai parameter, dan n adalah jumlah *Agent* [1].

$$V_1 = \frac{\sum (3 + 2 + 2 + 1 + 4 + 3 + 2 + 2 + 1 + 3)}{10}$$

2. Mencari Nilai Standar Deviasi Tiap Parameter

$$std(v3) = \sqrt{\frac{0.01 + 0.81 + 0.81 + 1.21 + 1.21 + 0.01 + 0.01 + 0.81 + 0.01 + 0.01}{10 - 1}}$$

Kemudian dilanjutkan dengan Kategori V3. Setelah itu akan didapatkan data baru yang telah distandarkan. Data dapat dilihat pada Tabel 3.

Tabel 3. Data Bibit Sawit Standar

No	V1	V2	V3
1	0.74	-0.51	-0.14
2	-0.316	0.5145	1.2197
3	-0.316	1.5435	1.2197
4	-1.37	-0.514	-1.4908
5	1.792	-0.514	-1.4908
6	0.7379	-0.514	-0.1355
7	-0.316	0.5145	-0.1355
8	-0.316	1.5435	1.2197
9	-1.37	-0.514	-0.1355
10	0.7379	-1.543	-0.1355

Kemudian dilakukan perhitungan menggunakan *Euclidean Distance*.

$$d_{rs} = \sum_{i=0}^k (x_{ri} - x_{si})^2 \quad (2)$$

Pada akhir perhitungan didapatkan matriks jarak yang baru yang menjadi sebuah *cluster* antara I dan O seperti berikut.

Tabel 4. Cluster 1

	AF	B	C	D	E	G	H	I	J
AF	0	1.9886	1.9977	1.7980	1.5704	1.0764	1.9977	2.1081	1.0289
B	1.9886	0	1.4552	3.0765	3.0765	1.3553	1.4552	1.9886	3.2105
C	1.9977	1.4552	0	3.0765	3.5628	1.7016	0	2.6801	3.5323
D	1.7980	3.0765	3.5628	0	3.1623	2.0017	3.5628	1.3553	2.7092
E	1.5704	3.0765	3.5628	3.1623	0	2.7092	4.0033	3.4405	2.0017
G	1.0764	1.3553	1.7016	2.0017	2.7092	0	1.7016	1.4731	2.3122
H	1.9977	1.4552	0	3.5628	4.0033	1.7016	0	2.0017	3.9763
I	2.1081	1.9886	2.6801	1.3553	3.4405	1.4731	2.0017	0	2.3459
J	1.0289	3.2105	3.5323	2.7092	2.0017	2.3122	3.9763	2.3459	0

Tabel 4 diperoleh dari mencari nilai terkecil dari matriks jarak. Dimana A dan F mempunyai nilai terkecil, yaitu 0 maka obyek A, C, F dan H bergabung menjadi satu *cluster*. Dilanjutkan dengan menghitung jarak antar *cluster* A dan F dengan obyek lain. Menggunakan :

$$\begin{aligned} d(A-F)B &= \min \{dAB, dFB\} \\ d(A-F)C &= \min \{dAC, dFC\} \end{aligned}$$

Kemudian mencari nilai terkecil dari matriks jarak, obyek AFJGID dan E mempunyai nilai terkecil yaitu 1.5704 maka obyek AFJGID dan E bergabung menjadi satu *cluster*.

Tabel 5. Cluster 8

	AFJGIDE	CHB
AFJGIDE	0	1.3553
CHB	1.3553	0

Tabel 5 merupakan nilai terkecil dari matriks jarak, obyek AFJGID dan E mempunyai nilai terkecil yaitu 1.5704 maka obyek AFJGID dan E bergabung menjadi satu *cluster*. Menggunakan :

$$\begin{aligned}
 D(\text{AFJGID-E})\text{CHB} &= \min \{d\text{AFJGIDCHB}, d\text{ECHB}\} \\
 &= \min \{1.3553, 3.0765\} \\
 &= 1.3553
 \end{aligned}$$

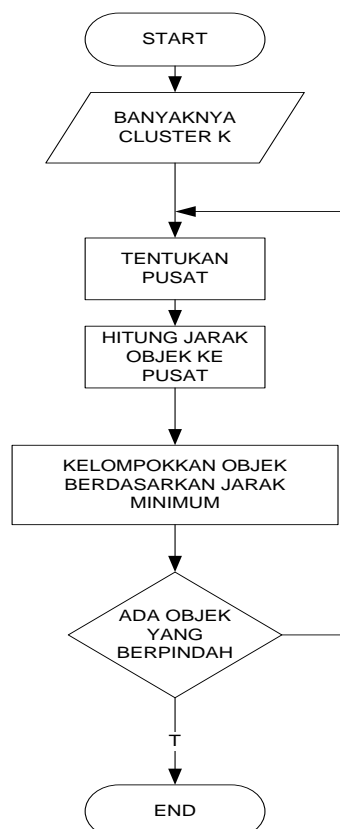
Sehingga didapatkan hasil akhir seperti yang ada di Tabel 6. Diperoleh Cluster bibit yang layak untuk dipergunakan.

Tabel 6. Cluster

Cluster	Bibit	Keterangan
0	AFJGIDE	Liar
1	CHB	Unggul

3.4 Pemodelan Sistem

Merupakan gambaran dari sistem yang akan dibangun. Dimodelkan dengan *flowchart*, seperti di Gambar 1.



Gambar 1. Flowchart Sistem

3.5 Hasil

Hasil dari sistem penerapan data mining untuk pengelompokan bibit sawit yang dibangun dapat dilihat pada gambar-gambar dibawah ini.

1) Form Menu Utama

Menu Utama merupakan tampilan awal pada saat aplikasi dijalankan. Tampilan menu utama dapat dilihat pada Gambar 2.



Gambar 2. Menu Utama

2) Form Tampilan Bibit Sawit

Tampilan Data bibit sawit merupakan halaman untuk memasukkan data bibit sawit. Bentuknya dapat dilihat pada Gambar 3.

KodeBibit	NamaBibit	Varietas	WilayahPem...
0001	Hibrida F-1 h...	D x P PPKS 5...	Aek Pancur, ...
0002	Hibrida F-1 h...	D x P PPKS 5...	Bukit Sentan...
0003	Hibrida F-1 h...	D x P PPKS 5...	PPKS- Unit M...
0004	Hibrida F-1 h...	D x P PPKS 5...	Dalu-Dalu, D...
0005	Hibrida F-1 h...	D x P PPKS 5...	Kalianta, Des...
0006	Hibrida F-1 h...	D x P PPKS 5...	Kalianta, Des...
0007	Hibrida F-1 h...	D x P PPKS 5...	Kalianta, Des...
0008	Hibrida F-1 h...	D x P PPKS 5...	Kalianta, Des...
0009	Hibrida F-1 h...	D x P PPKS 5...	Kalianta, Des...
0010	Hibrida F-1 h...	D x P PPKS 5...	Kalianta, Des...

Gambar 3. Tampilan Bibit Sawit

3) Form Hasil Perhitungan

Bentuk tampilan hasil perhitungan terdapat pada Gambar 4.

NoAnalisis	V1	V2	V3
A01	3	2	2
A02	2	3	3
A03	2	4	3
A04	1	2	1
A05	4	2	1
A06	3	2	2
A07	2	3	2
A08	2	4	3
A09	1	2	2
A10	3	1	2

NoAnalisis	V1	V2	V3
A01	0.777777...	-0.52941...	-0.18367...
A02	-0.33333...	0.52941...	1.653061...
A03	-0.33333...	1.588235...	1.653061...
A04	-1.44444...	-0.52941...	-2.02040...
A05	1.88888...	-0.52941...	-2.02040...
A06	0.777777...	-0.52941...	-0.18367...
A07	-0.33333...	0.52941...	-0.18367...
A08	-0.33333...	1.588235...	1.653061...
A09	-1.44444...	-0.52941...	-0.18367...
A10	0.777777...	-1.58823...	-0.18367...

V1	V2	V3
2.300000000000	2.500000000000	2.100000000000

V1	V2	V3
0.900000000000	0.944444444444	0.544444444444

NoAna...	KodeB...	Varietas	Wilaya...	Ketera...
A01	0001	D x P ...	Aek P...	LIAR
A02	0002	D x P ...	Bukit ...	UNGGUL
A03	0003	D x P ...	PPKS- ...	UNGGUL
A04	0004	D x P ...	Dalu...	LIAR
A05	0005	D x P ...	Kaliant...	LIAR
A06	0006	D x P ...	Kaliant...	LIAR
A07	0007	D x P ...	Kaliant...	LIAR
A08	0008	D x P ...	Kaliant...	UNGGUL
A09	0009	D x P ...	Kaliant...	LIAR
A10	0010	D x P ...	Kaliant...	LIAR

Gambar 4. Hasil Perhitungan

4. SIMPULAN DAN SARAN

4.1 Simpulan

Berdasarkan hasil penelitian dan pembahasan sebelumnya, maka penulis akan menarik kesimpulan dari penguraian tersebut :

1. Dari pengujian terhadap 20 data bibit sawit, menunjukkan metode *Single Linkage* mampu melakukan penentuan kelompok bibit sawit dengan jarak terpendek.
2. Penggabungan menggunakan metode *Single Linkage* menghasilkan total waktu rata-rata 4 detik dibandingkan metode *Complete Linkage* dan *Average Linkage* dengan *range* nilai antara 0 - 10.
3. Akurasi *clustering* sangat jelas dengan tingkat keakuratan mencapai 90 %. Dipengaruhi kesamaan pola perilaku bibit sawit yang digali melalui *cluster-cluster* 20 data bibit sawit yang terbentuk.

4.2 Saran

Penulis ingin memberikan sebuah saran dengan harapan dapat memberikan masukan bagi pengguna(*user*) dalam menggunakan sistem ini, yaitu :

1. Pada perancangan penerapan *data mining* untuk pengelompokan bibit sawit menggunakan Metode *Clustering* dengan Teknik *Single Linkage* agar lebih akurat dalam perhitungan dari nilai kriteria sebaiknya data diambil dari tempat/wilayah pembibitan sawit
2. Dalam pengelompokan bibit sawit sebaiknya dapat dikembangkan lebih jauh dengan aplikasi-aplikasi lainnya seperti Matlab, program berbasis *online*, dan aplikasi yang dapat mendeteksi bibit sawit tersebut unggul atau tidak sehingga dapat digunakan sebagai salah satu gambaran dalam pemilihan bibit sawit.

5. DAFTAR RUJUKAN

- [1] Iyung, P., 2015. *Panduan Teknis Budidaya Kelapa Sawit Untuk Praktisi Perkebunan*. (1) 1rd PT. Penebar Swadaya
- [2] Maruli, P., 2015. *Kupas tuntas Agribisnis kelapa sawit*. (2) 2rd, Prenada Media
- [3] Kusri, 2015. *Algoritma Data Mining*. (3) 1rd, Andi Offset
- [4] PM Booma., 2016. *Journal of Theoretical and Applied Information Technology*: Department of Computer and Engineering SRM University. *Classification Of Genes For Disease Identification Using Data Mining Techniques*, (4), pp.399-414
- [5] Jyoti Sinha., 2016. *International Journal of Science Technology & Engineering*: Department of Computer Science & Engineering DIMAT, Raipur. *Clustering Based Approach in Data Mining: A Review*, 42 (5), pp.363-368
- [6] SR Pande., 2012. *International Journal of Advanced Research in Computer and Communication Engineering* : Department of Computer Science SSES Amti's Science College. *Data Clustering Using Data Mining Techniques*, 42 (6), pp.494-499.
- [7] Sankar RG., 2011. *International Journal of Database Management Systems* : Enterprise DW/BI Consultant Tata Consultancy Services. *Customer Data Clustering Using Data Mining Technique*, 42 (7), pp.1-11.
- [8] T Sajana., 2016. *Indian Journal of Science and Technology* : KL University, Vaddeswaram. *A Survey on Clustering Techniques for Big Data Mining*, 42 (8), pp.1-12.
- [9] K.Kameshwaran., 2014. *International Journal of Computer Science and Information Technologies* : Department Of Computer Science & Engineering, Coimbatore Institute of Technology Coimbatore, Tamil Nadu, India. *Survey on Clustering Techniques in Data Mining*, 42 (9), pp.2272-2276.
- [10] K. Chitra., 2017. *International Journal of Computer Science and Mobile Computing* : School of Computer Studies, Rathnavel Subramaniam College of Arts and Science. *A Comparative Study of Various Clustering Algorithms in Data Mining*, 42 (10), pp.109-115.
- [11] Meenu Sharma., 2014. *International Journal Of Core Engineering & Management*: Department of computer science, SD Bansal college of technology. *Clustering In Data Mining : A Brief Review*, 42 (11), pp.117-126.
- [12] Z. Abdullah., 2015. *International Journal of Computer and Information Engineering* : Information Techlogy. *Hierarchical Clustering Algorithms in Data Mining*, 42 (12), pp.2201-2206.
- [13] R. Indhu., 2018. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*: Department of Computer Science, Bharathiar University, Coimbatore, Tamilnadu, India. *Comparison of Clustering Algorithm*, 42 (13), pp.218-223.