

PENERAPAN DATA MINING UNTUK CLUSTERING PEMINAT LAYANAN ICONNET BERDASARKAN WILAYAH AREA BALI MENGGUNAKAN METODE K-MEANS

Ida Bagus Suryadharma Santika¹⁾ Ketut Queena Fredlina²⁾ Putu Trisna Hady Permana³⁾

Program Studi Teknik Informatika¹⁾²⁾³⁾

STMIK Primakara, Denpasar, Bali¹⁾²⁾³⁾

gussuryadharma56@gmail.com¹⁾ queena@primakara.ac.id²⁾ trisnapermana.putu@gmail.com³⁾

ABSTRACT

PT. Indonesia Comnets Plus (ICON+) is a subsidiary of PT PLN (Persero) which is a company in the field of telecommunications networks and information technology. Starting his work in the telecommunications network business, the internet service provider offered is named Stroomnet which has now changed to ICONNET on May 31, 2021. Various packages are offered according to the service needs of the internet in several districts in Bali. In doing marketing, it is necessary to carry out an effective and efficient strategy. Basically, a business strategy can provide business information such as market segmentation, target market, and approaches to marketing. The main data were obtained from research subjects, both through documentation and observation. Looking at the development of market demand for 1 year, this study uses a clustering process to determine potential new service users so as to make it easier for company management in planning service providers and appropriate promotional strategies. The processing uses data mining techniques and the K-Means algorithm. In this study, it is hoped that the results of the research can help the retail division in determining promotion strategies for prospective new service users.

Keywords: *ICONNET, Data Mining, K-Means Clustering.*

ABSTRAK

PT. Indonesia Comnets Plus (ICON+) adalah entitas anak PT PLN (Persero) yang merupakan perusahaan dibidang jaringan telekomunikasi dan teknologi informasi. Mengawali kiprahnya dalam bisnis jaringan telekomunikasi, adapun layanan internet provider yang ditawarkan diberi nama Stroomnet yang sekarang berubah menjadi ICONNET pada tanggal 31 Mei 2021. Berbagai paket yang ditawarkan sesuai dengan layanan kebutuhan internet di beberapa wilayah kabupaten di Bali. Dalam melakukan pemasaran perlunya melakukan strategi yang efektif dan efisien. Pada dasarnya, strategi suatu bisnis dapat memberikan informasi bisnis seperti segmentasi pasar, target pasar, dan pendekatan dalam pemasaran. Data utama diperoleh dari subjek penelitian, baik melalui dokumentasi, dan observasi. Melihat perkembangan permintaan pasar selama 1 tahun, pada penelitian ini menggunakan proses klasterisasi untuk menentukan calon pengguna layanan baru sehingga mempermudah pihak manajemen perusahaan dalam perencanaan penyedia layanan dan strategi promosi yang tepat. Dalam pengolahannya menggunakan teknik penerapan data mining dan algoritma K-Means. Pada penelitian ini diharapkan hasil dalam penelitian dapat membantu pihak divisi retail dalam menentukan strategi promosi dalam calon pengguna layanan baru.

Kata Kunci : *ICONNET, Data Mining, K-Means Clustering.*

PENDAHULUAN

PT Indonesia Comnets Plus (ICON+) mengawali kiprahnya dalam bisnis telekomunikasi dan teknologi informasi pada tanggal 3 Oktober 2000 yang dulunya bernama Stroomnet dan sekarang berubah menjadi ICONNET pada tanggal 31 Mei 2021. Awalnya perjalanan ICON+ adalah memberikan dukungan secara total terhadap kebutuhan jaringan telekomunikasi PT PLN (Persero) sebagai induk perusahaan. Namun suatu permasalahan yang sering dihadapi oleh perusahaan penyedia jasa ialah dalam hal strategi promosi yang tepat, serta cara menentukan wilayah yang memiliki banyak peminat. Pengolahan data pada pengguna layanan baru dibutuhkan untuk membantu pihak ICON+ yang menyediakan layanan internet dalam mengetahui informasi penting berupa pengetahuan baru (*knowledge discovery*). Salah satu fungsi yang bisa didapat dengan menerapkan pengolahan data mining adalah fungsi *clustering* yang bisa digunakan untuk melakukan analisis pola tertentu dengan bersumber pada data dalam jumlah besar, *Clustering* merupakan salah satu metode data mining yang bersifat tanpa arahan (*unsupervised*) yang memiliki arti bahwa karakteristik tiap klaster tidak ditentukan sebelumnya melainkan berdasarkan kemiripan atribut-atribut dari suatu kelompok atau *cluster*. [19] Dengan adanya pengelompokan data seperti ini, diharapkan bagian marketing dapat melakukan analisa data dalam menentukan pengguna layanan baru disuatu wilayah dengan metode algoritma K-Means di PT. ICON+.

TINJAUAN PUSTAKA

Pengertian ICON+

PT. Indonesia Comnets Plus (ICON+) didirikan pada tanggal 3 Oktober 2000. PT. ICON+ berkomitmen untuk menyediakan jaringan, layanan dan konten telekomunikasi, khususnya mendukung teknologi dan sistem informasi PT. PLN Persero dan publik. Sebagai anak perusahaan yg dimiliki sepenuhnya oleh PLN, pada awalnya ICON+ berfokus untuk melayani kebutuhan PLN akan jaringan telekomunikasi. Seiring

menggunakan kebutuhan industri akan jaringan telekomunikasi menggunakan taraf availability serta reliability yang konsisten, Perseroan melihat peluang baru untuk menyebarkan usahanya yaitu dengan mengkomersialkan kelebihan kapasitas jaringan telekomunikasi ketenagalistrikan serat optik milik PLN pada Jawa serta Bali. Sesuai pemikiran tersebut, ICON+ mulai menjalin kerjasama menggunakan aneka macam perusahaan, terutama yang aktivitas operasionalnya membutuhkan jaringan telekomunikasi yang ekstensif serta handal. sampai waktu ini Perseroan melayani lebih dari 1127 perusahaan di Indonesia, di industri-industri utama yaitu telekomunikasi, perbankan, keuangan, pemerintahan serta manufaktur.

Data Mining

Data mining memiliki pengertian lain yaitu knowledge discovery ataupun pattern recognition merupakan suatu istilah yang digunakan untuk mendapatkan pengetahuan yang tersembunyi dari kumpulan data yang berukuran sangat besar. Tujuan utama data mining adalah untuk menemukan, menggali, atau menambang pengetahuan dari data atau informasi yang kita miliki[20].

Clustering

Clustering merupakan suatu metode untuk mencari dan mengelompokkan data yang memiliki kemiripan karakteristik (similarity) antara satu data dengan data yang lain. Tujuan utama dari metode clustering adalah pengelompokan sejumlah data atau objek ke dalam cluster (group) sehingga dalam setiap cluster dapat berisi data yang semirip mungkin. Dalam clustering metode ini berusaha untuk menempatkan objek yang mirip (jaraknya dekat) dalam satu cluster dan membuat jarak antar cluster sejauh mungkin. Ini berarti objek dalam satu cluster sangat mirip satu dengan lain dan berbeda dengan objek dalam cluster-cluster yang lain. [17]

K-Means

Algoritma K-means adalah algoritma yang terbaik dalam algoritma partitional clustering dan yang paling sering

digunakan di antara algoritma clustering lainnya karena kesederhanaan dan efisiensinya. Kelebihan Algoritma K-means itu sendiri merupakan algoritma clustering yang paling sederhana dan umum, hal ini dikarenakan K-means mempunyai kemampuan mengelompokkan data dalam jumlah yang cukup besar dengan waktu komputasi yang relatif cepat dan efisien. Namun, K-means mempunyai kelemahan yang diakibatkan oleh penentuan pusat awal cluster [17].



Gambar 1. Flowchart Proses Algoritma K-Means

Euclidean Distance

Merupakan rumus jarak yang umum digunakan karena sangat mudah dipahami dan mendukung hasil perhitungan cluster. Euclidean Distance menciptakan perhitungan jarak terpendek antara dua titik yang diperhitungkan. Perhitungan dilakukan selama belum mendapat nilai pengelompokan yang konvergen. [15]

$$d(i, j) = \sqrt{(x_{1i} - x_{1j})^2 + (x_{2i} - x_{2j})^2 + \dots + (x_{ki} - x_{kj})^2}$$

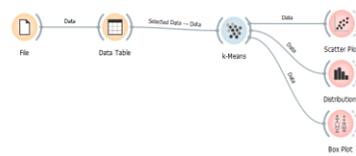
Keterangan:

- d (i, j) : jarak data ke i ke pusat cluster j
- x_{ki} : data ke i pada atribut data ke k
- x_{kj} : titik pusat ke j pada atribut ke k

Orange Data Mining

Data mining merupakan istilah pattern recognition merupakan algoritma untuk pengolahan data guna menemukan pola data menjadi pengetahuan baru [3]. Orange

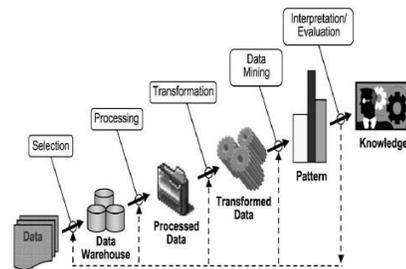
adalah teknologi pembelajaran mesin open source atau perangkat lunak penambangan data. Orange dapat digunakan untuk analisis dan visualisasi data eksploratif. Ini memberikan platform untuk pemilihan eksperimen, pemodelan prediktif, dan sistem rekomendasi dan dapat digunakan untuk penelitian genomik, biomedis, bioinformatika, dan pengajaran. Orange selalu disukai jika faktor inovasi, kualitas, atau keandalannya terlibat. Orange mempermudah pemakai bermain dengan data open source serta melaksanakan proses data analytics secara intuitif .[2]



Gambar 2. Pengolahan Program Orange

METODE PENELITIAN

Data mining adalah salah satu bagian proses Knowledge Discovery in Database (KDD) yang berfungsi untuk mengekstrak pola atau contoh berdasarkan data dengan menggunakan suatu algoritma yang spesifik.[16]



Gambar 3. Tahapan Proses KDD

1. Data Selection: pemilihan data berdasarkan sekumpulan data operasional perlu dilakukan sebelum tahap penggalian berita pada KDD dimulai.
2. Preprocessing: sebelum proses data mining dapat dilaksanakan, perlu dilakukan proses cleaning dengan tujuan untuk membuang duplikasi data, mempelajari data yang

inkonsisten, dan memperbaiki kesalahan dalam data, misalnya kesalahan cetak (tipografi). Juga dilakukan proses enrichment, yaitu proses “memperkaya” data yang telah ada dengan data atau berita lain yang relevan dan dibutuhkan untuk KDD, misalnya data atau informasi eksternal.

3. Transformation: yaitu proses coding dalam data yang sudah dipilih, sebagai akibatnya data tersebut sesuai untuk proses data mining. Proses coding pada KDD adalah proses kreatif dan sangat tergantung dalam jenis atau pola informasi yang akan dicari pada database.

4. Data mining: proses mencari pola atau informasi menarik pada data terpilih dengan memakai teknik atau metode eksklusif.

5. Interpretation/Evaluation: pola informasi yang didapatkan berdasarkan proses data mining perlu ditampilkan pada bentuk yang mudah dimengerti oleh pihak yang berkepentingan. Tahap ini adalah bagian dari proses KDD yang diklaim menggunakan interpretation. Tahap ini meliputi pemeriksaan apakah pola atau informasi yang ditemukan bertentangan menggunakan fakta atau hipotesa yang terdapat sebelumnya atau tidak.

Teknik Pengumpulan Data

Adapun jenis pengumpulan data yang digunakan dalam penelitian. Data penelitian ini berasal dari data yang dikumpulkan melalui dataset yang dikumpulkan oleh PT.ICON+ Divisi Retail tahun 2020-2021.

Analisis Kebutuhan Perangkat Keras

Perangkat keras yang dibutuhkan dalam Penerapan Data Mining Untuk Clustering Peminat Layanan Iconnet Berdasarkan Wilayah Area Bali Menggunakan Metode K-Means, terdiri dari:

- Processor AMD A9
- Memory 256 GB.
- Harddisk 1 GB.
- Mouse dan keyboard sebagai piranti input.
- Printer sebagai piranti output.

Analisis Kebutuhan Perangkat Lunak

Perangkat lunak yang dibutuhkan dalam Penerapan Data Mining Untuk Clustering Peminat Layanan Iconnet Berdasarkan Wilayah Area Bali Menggunakan Metode K-Means, terdiri dari:

- Microsoft Windows 11
- *Web browser* (Chrome, Mozilla Firefox, Opera).
- *Orange Data Mining*

Jenis dan Sumber Data

Jenis data yang digunakan adalah data kuantitatif dan data kualitatif. Data kuantitatif adalah jenis data yang dapat dihitung, berupa angka atau nominal. Lebih spesifik lagi, data yang digunakan berupa data matriks, yaitu jenis data yang memiliki iobjek dan atribut [8]. Data kualitatif adalah jenis data yang berupa informasi data diri seperti nama, jenis kelamin, ID pelanggan, dst. Sumber Data yang digunakan dalam penelitian adalah Data Primer dan Data Sekunder. Sumber data primer merupakan sumber data yang diperoleh secara langsung dari sumber asli dan tidak melalui media perantara.

Data historis transaksi penjualan yang digunakan diperoleh secara langsung dari objek penelitian melalui dokumentasi. Sedangkan data sekunder merupakan sumber data penelitian yang diperoleh secara tidak langsung melalui media perantara diperoleh dan dicatat oleh pihak lain. Data sekunder pada umumnya berupa bukti catatan atau laporan historis yang dipublikasikan. Data sekunder yang peneliti maksud dalam penelitian ini adalah sumber data yang digunakan untuk menunjang kelengkapan teori data primer.[8]

Tepat & Waktu Penelitian

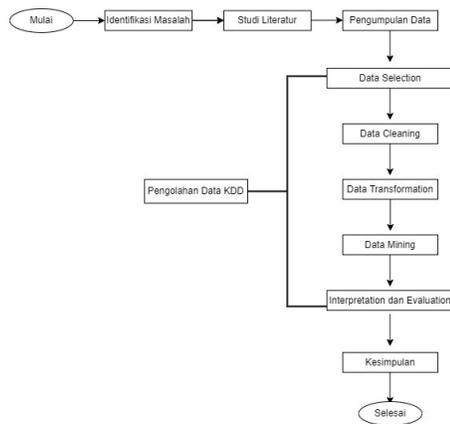
Penelitian ini dilakukan di PT Indonesia Comnets Plus (ICON+) yang beralamat di Gunung Mandali Wangi No 15. Waktu penelitian dimulai dari bulan September 2021 sampai dengan Maret 2022.

Metode Pengumpulan Data

Berdasarkan sumber data yang digunakan pada penelitian proyek akhir ini, maka metode pengumpulan data yang penulis gunakan adalah sebagai berikut:

1. Dokumentasi, adalah metode ipengumpulan data dengan mengumpulkan catatan-catatan atau dokumen-dokumen. Data yang telah penulis dapatkan dari metode pengumpulan data historis yang telah dicatat oleh pihak lain, penulis kumpulkan menjadi satu sehingga menjadi sebuah dokumen yang siap digunakan untuk kepentingan penelitian.
2. Studi Pustaka, adalah mempelajari karya ilmiah, buku ilmiah, dan sumber ilmiah lainnya yang sesuai dengan penelitian dan memiliki hubungan dengan masalah yang diteliti. Referensi ilmiah yang penulis gunakan adalah sumber-sumber yang terdapat dalam daftar kepustakaan.[8]

Alur Penelitian



Gambar 4. Alur Penelitian

Variabel Penelitian

Variabel Penelitian yang digunakan dalam penelitian ini ialah data peminat layanan ICONNET dari PT. ICON+ Regional Bali dan Nusa Tenggara antara lain umur, kabupaten/kota, desa/kecamatan, jenis kelamin, jenis tempat tinggal.

Jadwal Penelitian

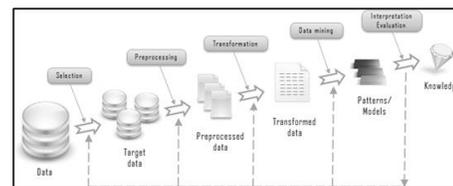
No	Kegiatan	Sept.	Okt.	Nov.	Des.	Jan.	Feb.	Mar.	Apr.	Mei.
1	Identifikasi Masalah									
2	Studi Literatur									
3	Pengumpulan Data									
4	Penyusunan Proposal Skripsi									
5	Penyusunan Skripsi									
6	Pengolahan data dengan KDD									
7	Implementasi Orange									
8	Kesimpulan									

Gambar 5

HASIL & PEMBAHASAN

Penelitian diawali dengan melakukan observasi langsung ke PT. ICON+ SBU Bali Nusra untuk menggali informasi menggali informasi data peminat layanan ICONNET pelanggan 2020/2021. Data hasil pelanggan terdiri dari kode jenis kelamin, umur, kota/kabupaten, kecamatan, status tempat tinggal, status presales. Dataset yang diberikan oleh bagian Retail dalam bentuk Excel yaitu sebagai database yang terdiri dari 14791 data dari tahun 2020-2021.

Penerapan Knowledge Discovery from Database (KDD)



Gambar 5. Proses Tahapan KDD

Data Selection

Tahap pertama untuk melakukan clustering data yaitu melakukan data selection. Dari dataset dari calon pelanggan layanan ICONNET dari yang berjumlah 14791 data mejadi 3407 data. Pada tahapan ini juga dilakukannya seleksi atribut-atribut data yang dianggap relevan dan layak untuk dipakai. Dalam penelitian ini, atribut yang digunakan ada 6 (enam) yaitu jenis kelamin, umur, kota/kabupaten, kecamatan, status tempat

tinggal, status presales. Lalu pada atribut ktp dirubah menjadi atribut umur agar mempermudah proses pengolahan data.

NO	Email	Nama Lengkap	ID Pelanggan	Nomor Handphone	Nomor KTP
1	Eka*****	I Gusti*****	14399081***	89516385***	510*****
2	Nyoman*****	Nyoman*****	32195371***	87763308***	510*****
3	Putu*****	Putu*****	5514002171***	87338094***	510*****
4	Thawan*****	Kurniawan*****	1118612***	87757019***	527*****
5	Indah*****	Indah*****	441100945***	8190714***	520*****
6	Net*****	I Putu*****	32004775***	81246413***	510*****
7	Net*****	I Putu*****	86000396***	81804515***	510*****
8	Kusuma*****	Ni Putu*****	441500030***	87861520***	527*****
9	Bumdes*****	I Made*****	551810041***	8170667***	510*****
10	Bumdes*****	Gideon****	32900008***	8170667***	510*****

Gambar 6. Contoh Dataset Pelanggan Yang Mendaftar

Pada tahap preprocessing dilakukan pembersihan data. Data yang tidak memiliki nilai maka tidak akan digunakan dalam dataset. Untuk mendapatkan data yang valid dilakukan tahapan cleaning. Proses pembersihan data dilakukan pada Microsoft Excel. meliputi, e-mail, nama lengkap, ID pelanggan, nomor handphone, provinsi, alamat tempat tinggal, rencana tambah daya, pilihan paket, komentar atau saran, dan titik koordinat pemasangan layanan sehingga dilakukan proses pembersihan data untuk data yang berkualitas.

NO	Nama lengkap	Nomor PLN	Nomor HP	KTP
1	I Made*****	86055339***	+628814706***	510*****
2	I Wayan*****	-	08123949***	510*****
3	*****	551800239***	081936297***	510*****
4	I Dewa*****	-	081936297***	510*****
5	Iulianto*****	-	-	510*****
6	I Ketut*****	86030065***	08123801***	510*****
7	I Made*****	-	-	510*****
8	I Wayan*****	-	-	510*****
9	I Gede Putu*****	184232***	-	510*****
10	Ibelia Maizar*****	14345587***	08573757***	510*****

Gambar 7. Contoh Data yang Tidak Digunakan

Pada dataset tersebut terdapat beberapa atribut yang tidak diperlukan yaitu, e-mail, nama lengkap, ID pelanggan, nomor handphone, alamat tempat tinggal, rencana tambah daya, pilihan paket, komentar atau saran, dan titik koordinat pemasangan layanan sehingga dilakukan proses pembersihan data untuk data yang melakukan pengolahan data pada Orange Data Mining. Pada atribut ktp dirubah menjadi atribut umur agar mempermudah proses pengolahan data pada tools orange, sehingga diperoleh jumlah dalam dataset 3407 data yang selanjutnya akan diolah pada aplikasi orange untuk mendapatkan hasil yang diinginkan.

Transformation

Merupakan proses transformasi pada data yang telah dipilih, sehingga data tersebut sesuai untuk proses data mining. Proses ini merupakan proses kreatif dan sangat tergantung pada jenis atau pola informasi yang akan dicari dalam basis data. Lalu tahap ini dilakukan penyesuaian data ke dalam bentuk .csv agar data dapat diolah dengan menggunakan tools data mining dengan mengimplementasikan algoritma K-Means.

NO	Jenis Kelamin	Umur	Kota/Kabupaten	Kecamatan	Status Tempat Tinggal	Status Presales
1	Pria	23	Badung	Mengen	Milik Sendiri	Tercover
2	Pria	49	Bangli	Bangli	Milik Sendiri	Tercover
3	Wanita	36	Tabanan	Selamadag Timur	Milik Sendiri	Tercover
4	Pria	43	Tabanan	Selamadag Timur	Milik Sendiri	Tidak Tercover
5	Pria	28	Tabanan	Selamadag Timur	Milik Sendiri	Tidak Tercover
6	Pria	51	Bangli	Bangli	Milik Sendiri	Tercover
7	Pria	18	Bangli	Bangli	Milik Sendiri	Tercover
8	Pria	42	Buleleng	Buleleng	Milik Sendiri	Tidak Tercover
9	Wanita	51	Garuda	Garuda	Milik Sendiri	Tidak Tercover
10	Pria	20	Bangli	Bangli	Milik Sendiri	Tercover

Gambar 8. Contoh Data Transformation

Data Mining

Pada tahap ini dilakukan pengclusteran dengan menggunakan algoritma k-means yang dibangun sesuai tujuan penelitian. Pengujian dataset untuk mengclustering peminat layanan ICONNET yang menggunakan dalam 3 cluster yaitu sangat berminat, berminat, dan tidak berminat yang diimplementasikan dengan aplikasi orange.

Evaluation

Pada tahap ini adalah fase lanjutan. Evaluasi dilakukan secara mendalam dengan tujuan agar hasil pada tahap pemodelan sesuai dengan target yang ingin dicapai dalam tahap business understanding.

Perhitungan Manual Algoritma K-Means Menggunakan Microsoft Exel

Berikut contoh perhitungan clustering dari 10 data pelanggan peminat layanan ICONNET, yang diambil dari data acak yang sudah di transformasi sesuai dengan kebutuhan penelitian seperti pada tabel 4.4 di bawah ini:

NO	Jenis Kelamin	Umur	Kota/Kabupaten	Kecamatan	Status Tempat Tinggal	Status Presales
1	Pria	34	Tabanan	Kediri	Milik Sendiri	Tidak Tercover
2	Pria	40	Tabanan	Selemadeg	Milik Sendiri	Tercover
3	Pria	42	Tabanan	Selemadeg	Milik Sendiri	Tidak Tercover
4	Pria	47	Denpasar	Denpasar Utara	Milik Sendiri	Tercover
5	Wanita	31	Bangli	Tembuku	Milik Sendiri	Tidak Tercover
6	Pria	29	Gianyar	Tegallalang	Milik Sendiri	Tidak Tercover
7	Pria	34	Denpasar	Denpasar Barat	Milik Sendiri	Tercover
8	Pria	49	Tabanan	Penebel	Milik Sendiri	Tidak Tercover
9	Pria	34	Denpasar	Denpasar Barat	Milik Sendiri	Tercover
10	Pria	49	Tabanan	Penebel	Milik Sendiri	Tidak Tercover

Gambar 9. Dataset Percobaan

Titik Pusat Awal Cluster

No Data	Jenis Kelamin	Umur	Kota/Kabupaten	Kecamatan	Status Tempat Tinggal	Status Presales
2	2	40	4	42	1	1
7	2	34	4	41	1	2
5	1	31	1	11	2	2

Gambar 10. Titik Pusat Awal Cluster

Menghitung jarak setiap data ke pusat cluster awal menggunakan persamaan *Euclidean Distance* sebagai berikut:

$$d(i, j) = \sqrt{(x_{1i} - x_{1j})^2 + (x_{2i} - x_{2j})^2 + \dots + (x_{ki} - x_{kj})^2}$$

Gambar 11, Rumus *Euclidean Distance*

$$d(1,2) = \sqrt{(2-2)^2 + (34-30)^2 + (4-4)^2 + (41-42)^2 + (1-1)^2 + (2-1)^2} = 38$$

Gambar 12. Cara menghitung dengan *Euclidean Distance*

Berdasarkan hasil ketiga perhitungan diatas dapat disimpulkan bahwa jarak data pertama yang paling dekat adalah cluster-2, sehingga data pertama dimasukkan ke dalam cluster-2.

NO	C1	C2	C3	Jarak Terpendek
1	38	0	920	0
2	0	38	1054	0
3	5	65	1093	5
4	494	574	360	360
5	1054	920	0	0
6	244	126	410	126
7	440	366	134	134
8	83	229	1359	83
9	440	366	134	134
10	83	229	1359	83

Gambar 13. Hasil Perhitungan *Cluster* Iterasi 1

No.	C1	C2	C3
1		1	
2	1		
3	1		
4			1
5			1
6		1	
7			1
8	1		
9			1
10	1		

Gambar 14. Hasil Pengelompokan Iterasi 1

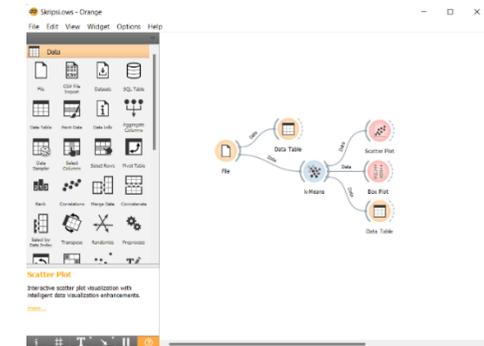
Setelah ditempatkan pada cluster paling dekat, selanjutnya pusat cluster baru dihitung kembali berdasarkan nilai rata-rata data yang diinisialisasi sebelumnya. Pada perhitungan iterasi ketiga dan keempat tidak mengalami perubahan, maka proses iterasi selesai dan diperoleh 3 cluster dengan 4 iterasi. Berikut tabel hasil akhir dari analisa perhitungan data peminat layanan ICONNET dengan menggunakan algoritma k-means.

Cluster	Hasil Akhir	
	Centroid Akhir	Jumlah Responden
1	5	1
2	1,2,3,4,6,7,8,9	9
3	0	0

Gambar 15. Hasil Analisa

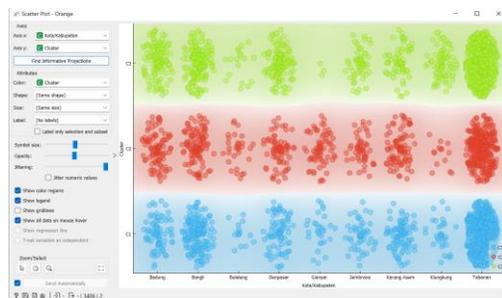
Pada hasil perhitungan manual diatas dapat disimpulkan secara sementara pada cluster 1 hanya ada 1 data yaitu dengan karakteristik berjenis kelamin wanita dengan umur 31 tahun berasal dari kabupaten Bangli di daerah Tembuku lalu dengan status tempat tinggal milik sendiri dan wilayah yang belum tercover oleh jaringan. Lalu pada cluster 2 ada 9 data yang bisa dirata-ratakan adalah berjenis kelamin pria dengan umur 29-49 tahun berasal dari kabupaten Tabanan di daerah Selemadeg dan Penebel lalu dengan status tempat tinggal milik sendiri dan wilayah yang sudah tercover oleh jaringan. Dan yang terakhir pada cluster 3 tidak ada data yang tersedia.

Implementasi Aplikasi Orange



Gambar 16. Rancangan Model K-Means

Pengujian dengan *software orange* membantu untuk proses *clustering* menjadi lebih mudah, efektif dan efisien. *Orange* selalu disukai jika faktor inovasi, kualitas, atau keandalannya terlibat. *Orange* mempermudah pemakai bermain dengan *data open source* serta melaksanakan proses *data analytics* secara intuitif [2].



Gambar 17. Tampilan dari fitur *Scatter Plot*

Widget scatter plot adalah gambaran yang menunjukkan kemungkinan hubungan (korelasi) antara pasangan dua macam variabel dan menunjukkan keeratan hubungan antara dua variabel tersebut yang sering diwujudkan sebagai koefisien korelasi.



Gambar 18. Tampilan dari setiap Kabupaten

Box Plot juga berguna untuk menemukan properti dari kumpulan data tertentu, misalnya, satu set *instance* yang didefinisikan secara manual di *widget* lain. Hasil dari perhitungan pada *tools orange* tersebut menunjukkan pada C1 terdapat 1166 data, lalu pada C2 terdapat 917 data dan yang terakhir C3 terdapat 1323 data, terlihat pada gambar 4.18 dibawah ini.

Berdasarkan hasil grafik *Box Plot* yang ditunjukkan pada Gambar 4.18 dapat diketahui bahwa pada setiap cluster memiliki beberapa perbedaan setiap karakteristiknya yang dimana sebagai berikut:

- Pada C1 memiliki tingkat peminat yang sedang dengan karakteristik rata-rata berjenis kelamin pria dengan umur 46-55 tahun di kabupaten Karangasem dengan status tempat tinggal milik sendiri lalu yang terakhir status presales tercover.
- Pada C2 memiliki tingkat peminat yang rendah dengan karakteristik rata-rata berjenis kelamin pria dengan umur 25-38 tahun di kabupaten Buleleng dan Gianyar dengan status tempat tinggal milik sendiri lalu yang terakhir status presales tidak tercover.
- Pada C3 memiliki tingkat peminat yang tinggi dengan karakteristik rata-rata berjenis kelamin pria dengan umur 24-36 tahun di kabupaten Bangli, Badung, Denpasar, Jembrana, Klungkung, Tabanan dengan status tempat tinggal milik sendiri lalu yang terakhir status presales tercover.

SIMPULAN

Setelah dilakukan pengelompokan data pengguna layanan ICONNET dapat disimpulkan dengan melihat hasil pada setiap cluster yang dimana pada C3 memiliki jumlah peminat yang besar sebanyak 1323 data peminat dengan karakteristik rata-rata berjenis kelamin pria dengan umur 24-36 tahun di kabupaten Bangli, Badung, Denpasar, Jembrana, Klungkung, Tabanan dengan status tempat tinggal milik sendiri lalu yang terakhir status presales tercover.

DAFTAR PUSTAKA

- [1] Eska, "Penerapan Data Mining untuk prediksi penjualan wallpaper menggunakan Algoritma C4.5" JURTEKSI, vol. 2, No. 2, 2016
- [2] Abdillah, G., Putra, F. A., Renaldi, F., Informatika, P. S., Jenderal, U., & Yani, A. (2016). Penerapan Data Mining Pemakaian Air Pelanggan Untuk Menentukan iKlasifikasi Potensi Pemakaian Air Pelanggan Baru Di Pdam Tirta Raharja Menggunakan Algoritma K-Means. In Seminar Nasional Teknologi Informasi dan Komunikasi (pp. 498-506).
- [3] S. iAgustina, D. Yhudo, H. Santoso, N. Marnasusanto, A. Tirtana dan F. Khusnu, "Clustering Kualitas Beras Berdasarkan Ciri Fisik Menggunakan Metode K-Means," Universitas Brawijaya Malang, Malang, 2012
- [4] Budiman, R. (2019). Penerapan Data Mining Untuk Menentukan Lokasi Promosi Penerimaan Mahasiswa Baru Pada Universitas Banten Jaya (Metode K-Means Clustering). ProTekInfo (Pengembangan Riset dan Observasi Teknik iInformatika), 6(1), 6-14.
- [5] Hastuti, Noor Fitriana. "Pemanfaatan metode K-Means Clustering dalam penentuan penerima beasiswa." (2013).
- [6] Merliana, Ernawati, Santoso, "Analisa Penentuan Jumlah Cluster Terbaik Pada Metode K-Means Clustering." Prosiding seminar nasional multi disiplin ilmu & call for papers unisbank, 2016.
- [7] Ediyanto, Mara, Satyahadewi, "Pengklasifikasian Karakteristik Dengan Metode K-Means Cluster Analysis." Buletin Ilmiah Mat. Stat. dan Terapannya, vol. 02, No. 2, pp. 133 – 136, 2013.
- [8] Muningsih, E., & Kiswati, S. (2015). Penerapan metode K-Means untuk Clustering produk online shop dalam penentuan stok barang. *Bianglala Informatika*, 3(1).
- [9] Utomo, D. P., & Mesran, M. (2020). Analisis Komparasi Metode Klasifikasi Data Mining dan Reduksi Atribut Pada Data Set Penyakit Jantung. *Jurnal Media Informatika Budidarma*, 4(2), 437-444.
- [10] Setiawan, R. (2017). Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Untuk Menentukan Strategi Promosi Mahasiswa Baru (Studi Kasus: Politeknik Lp3i Jakarta). *Jurnal Lentera Ict*, 3(1), 76-92.
- [11] Handayanto, A., Latifa, K., Saputro, N. D., & Waliansyah, R. R. (2019). Analisis dan Penerapan Algoritma Support Vector Machine (SVM) dalam Data Mining untuk Menunjang Strategi Promosi. *JUITA: Jurnal Informatika*, 7(2), 71-79.
- [12] Tonggiroh, M., i& Jufri, M. T. (2018, October). Data Mining Strategi Promosi Pada Universitas Yapis Papua Menggunakan Algoritma K-Means Clustering. In Seminar Nasional Teknologi Informasi dan Komunikasi (SEMNASITIK) (Vol. 1, No. 1, pp. 587-594).
- [13] Kusumo, H., Sedyono, E., & Marwata, M. (2019). Analisis algoritma apriori untuk mendukung strategi promosi perguruan tinggi. *WJIT: Walisongo Journal of Information Technology*, 1(1), 51-62.
- [14] Nurhafida, S. I., & Sembiring, F. (2021, September). ANALISIS TEXT CLUSTERING MASYARAKAT DI TWITTER MENGENAI MCDONALD'SXBTS MENGGUNAKAN ORANGE DATA MINING. In Seminar Nasional Sistem Informasi dan Manajemen Informatika Universitas Nusa Putra (Vol. 1, No. 01, pp. 28-35).
- [15] Falentino Sembiring, Octaviana, Sudin Saepudin, "Implementasi Metode K-Means Dalam Pengklasteran Daerah Pungutan Liar Kabupaten Sukabumi (Studi Kasus: Dinas Kependudukan dan Pencatatan Sipil)", Vol.14, No.1, April i020.
- [16] Pormes, R., & Manongga, D. H. (2019). Penggunaan Algoritma Clustering K-Means Untuk Melihat Daerah-Daerah Penyuplai

Mahasiswa Di Universitas Kristen Satya Wacana, Salatiga. *Jurnal Teknik Informatika dan Sistem Informasi*, 5(3).

[17] Almadya, A., & Saepudin, S. (2021, September). PENERAPAN DATA MINING K-MEANS CLUSTERING UNTUK MENGELOMPOKKAN BERBAGAI JENIS MERK SMARTPHONE. In *Seminar Nasional Sistem Informasi dan Manajemen Informatika Universitas Nusa Putra* (Vol. 1, No. 01, pp. 241-249).

[18] Yunita, F. (2018). Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Pada Penerimaan Mahasiswa Baru. *Sistemasi: Jurnal Sistem Informasi*, 7(3), 238-249.

[19] A. Asroni, H. Fitri, and E. Prasetyo, "Penerapan Metode Clustering dengan Algoritma K-Means pada Pengelompokan Data Calon Mahasiswa Baru di Universitas Muhammadiyah Yogyakarta (Studi Kasus: Fakultas Kedokteran dan Ilmu Kesehatan, dan Fakultas Ilmu Sosial dan Ilmu Politik)," *Semesta Tek.*, vol. 21, no. 1, pp. 60–64, 2018, doi: 10.18196/st.211211.

[20] S. Rony, "Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Untuk Menentukan Strategi Promosi Mahasiswa Baru (Studi Kasus: Politeknik Lp3i Jakarta)," *J. Lentera Ict*, vol. 3, no. 1, pp. 76–92, 2016