

EVALUASI CLUSTER SOCIAL MEDIA DATA IN TOURISM DOMAIN MENGUNAKAN K-MEANS CLUSTERING

¹Rena Nainggolan, ²Fenina Adline Twince Tobing, ³Emma Rosinta Simarmata,
⁴Resianta Perangin-angin

^{1,3,4}Komputerisasi Akuntansi, Universitas Methodist Indonesia

²Teknik Informatika, Universitas Multimedia Nusantara

Email: renanain66olan@gmail.com

DOI: <https://doi.org/10.46880/jmika.Vol4No1.pp89-93>

ABSTRAK

Clustering adalah salah satu teknik Data Mining. *Clustering* bekerja dengan cara menggabungkan sejumlah data atau objek kedalam satu klaster, dengan tujuan setiap data dalam satu klaster tersebut akan mempunyai data yang semirip mungkin dan berbeda dengan data atau objek yang berada pada kelompok lain. K-Means Clustering memiliki kemampuan untuk melakukan komputasi yang relatif cepat dan efisien dalam menggabungkan data dalam jumlah yang cukup besar. Dalam penelitian ini, peneliti akan menggunakan metode K-mean clustering yang diterapkan pada data mining pada *Online Reviews* pada data TripAdvisor. Implementasi proses *K-Means Clustering* menggunakan Weka, Atribut yang digunakan adalah galeri seni, klub dansa, bar jus, restoran, museum, resor, taman atau tempat piknik, pantai, teater, dan lembaga keagamaan. Menghasilkan jumlah *cluster* 4 ($k=4$) dengan *cluster* pertama sebanyak 178 (18%) *reviews traveler*, *cluster* kedua 243 (25%) *reviews traveler*, *cluster* ketiga 228 (23%) *reviews traveler*, *cluster* keempat 331(34%) *reviews traveler*.

Kata Kunci: *Data Mining, K-Means Clustering, Cluster, Social Media, TripAdvisor.*

PENDAHULUAN

Perkembangan teknologi sangat pesat di era globalisasi ini, untuk mempermudah pekerjaan banyak aspek yang bisa dimanfaatkan, demikian juga untuk arus informasi. Dengan menerapkan teknologi komputer diberbagai bidang, misalnya pendidikan, hiburan, kesehatan, sektor pariwisata, kuliner dan lain sebagainya.

Terlebih pada bidang bisnis saat ini sangat menuntut pihak pebisnis untuk menggunakan teknologi komputer supaya bisa bersaing di era globalisasi saat ini, yang disebut e-commerce, Transaksi dalam dunia internet diartikan sebagai E-Commerce (Reddy&Janna, 2012) misalnya TripAdvisor, situs ini membantu para wisatawan yang ingin melakukan perjalanan baik dalam pemesanan hotel, atau kunjungan ke tempat wisata lainnya. Sehingga perjalanan wisatawan lebih tersusun atau terencana dengan baik.

Dengan kemudahan dan keunggulan dengan menggunakan E-commerce, perusahaan-perusahaan melakukan promosi secara online dengan menggunakan elektronik dan internet marketing dengan merancang website sendiri dan memasukkan informasi tentang fasilitas-fasilitas dan keunggulan yang dimiliki oleh perusahaannya ke web-web yang terkenal, dengan tujuan menjadikan daya tarik kepada calon wisatawan dengan harapan calon wisatawan

atau traveler tertarik untuk memilih dan memutuskan untuk menggunakannya.

Selain melakukan promosi produk dan fasilitas yang dimiliki setiap destinasi wisata, Departemen E-Commerce juga menangani review dari si pengguna E-commerce tersebut, bisa dilihat dari penggunaan internet kini telah menjadi alat komunikasi dalam kehidupan sehari-hari baik dalam mencari informasi atau kegunaan lainnya.

Dengan meningkatnya penggunaan internet, mendorong para pengelola hotel, restoran, club dance, dan berbagai destinasi wisata lainnya untuk lebih meningkatkan efektifitas promosi, salah satu layanan yaitu reservasi hotel atau destinasi wisata lainnya secara online yang sering disebut dengan istilah online booking, dengan adanya online booking maka muncullah website travel review yang membantu konsumen untuk dapat berinteraksi langsung dengan si penyedia jasa atau tempat wisata, konsumen bisa melakukan tawar menawar, memberikan saran, ataupun memberikan kesan ataupun feedback dalam penggunaan fasilitas yang telah disediakan, yang salah satunya adalah TripAdvisor.

TripAdvisor adalah salah satu situs perjalanan terbesar di dunia, yang menyediakan berbagai fasilitas sehingga rencana perjalanan wisatawan berjalan dengan baik dan menyenangkan. Pada tahun 2000 didirikan TripAdvisor di Needham,

Amerika Serikat di kota Massachusset. Pertama TripAdvisor fungsinya untuk mempertemukan produsen dengan produsen yang disebut sebuah situs B2B (Businnes to bussines), dengan berjalannya waktu situs ini dimanfaatkan oleh para konsumen khususnya traveler untuk menshare pengalaman selama melakukan perjalanan menggunakan fasilitas situs tersebut, sejak itulah TripAdvisor berubah menjadi review site untuk pariwisata menjadi salah satu media online untuk umum.

Untuk meminimalisasi dampak negatif, sebelum melakukan booking online traveler harus mencari informasi sebanyak mungkin tentang tujuan wisata yang akan dituju. Cara yang paling mudah adalah melihat peringkat dari ulasan atau rating pada destinasi wisata tersebut.

KAJIAN PUSTAKA

Data Mining

Ada dua istilah dalam data mining yaitu; seperti knowledge discover ataupun pattern recognition. Masing-masing mempunyai arti yang berbeda-beda dan memiliki ketetapan satu sama yang lain, Menurut (Agusta, 2007) tujuan data mining untuk memperoleh pengetahuan yang masih tersembunyi dalam bongkahan data, sedangkan menurut (Berry et al, 2006) adalah pengenalan pola dalam bongkahan data yang akan digali disebut pattern recognition

K-Mean Clustering

Metode K-Means clustering merupakan metode clustering yang dikenalkan oleh (Susanto et al, 2011). Menurut (Lee et al, 2014). K-means clustering adalah metode yang sederhana dan efisien.

Cara kerja Metode K-Means dengan mengelompokkan data atau objek yang memiliki karakter yang sangat mirip dalam dalam satu cluster/kelompok dan akan mengelompokkan data/objek ke cluster yang lain yang mempunyai karakteristik yang berbeda dan pada akhirnya akan menghasilkan suatu cluster atau kelompok yang memiliki tingkat kemiripan yang sangat tinggi. Langkah-langkah melakukan clustering dengan metode K-Means adalah sebagai berikut (Jiawei, 2011)

1. Tentukan K atau jumlah cluster
2. Pusat-pusat cluster diberi nilai awal dengan angka random. Tetapi dalam memberikan inialisasi k pusat cluster dapat dilakukan dengan berbagai cara, dan yang paling serind adalah secara random.
3. Dalam menempatkan objek berada di cluster mana ditempatkan, maka dihitung berdasarkan jarak antara dua objek, demikian juga dengan jarak

antara pusat cluster dengan objek. Untuk menghitung jarak semua data ke setiap titik pusat cluster dapat menggunakan teori jarak euclidean yang dirumuskan sebagai berikut:

$$D(i, j) = \sqrt{(X1i - X1j)^2 + (X2i - X2j)^2 + \dots + (Xki - Xkj)^2}$$

Dimana:

$D(i, j)$ = 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

4. Menghitung kembali jarak keanggotaan cluster sekarang dengan pusat cluster, rata-rata dari semua data/objek dalam cluster tertentu merupakan pusat cluster, dengan menggunakan mean (nilai rata-rata) atau median
5. Lakukan kembali pada setiap objek dengan cluster baru, jika pusat cluster baru tidak berganti atau berubah, maka proses clustering berhenti, jika tidak maka akan kembali ke tahap ketiga sampai pusat cluster tidak berubah

Online Customer Reviews

Menurut (Almana et, 2013) Salah satu promosi pada penjualan online dalam bentuk Word of Mouth Communication disebut Online Customer Reviews (OCRs). Konsumen akan memperoleh informasi mengenai manfaat dari produk yang ditawarkan, hal ini akan membantu konsumen untuk membandingkan dengan produk yang sama pada penjual online yang lainnya sehingga konsumen tidak tertipu atau dirugikan, hal ini akan banyak memberikan keuntungan terhadap konsumen tanpa harus mengunjungi satu persatu si penjual secara langsung (Chen et al, 2008)

Clustering

Cara kerja clustering adalah membentuk satu kelompok dari kumpulan berbagai objek fisik atau abstrak ke dalam satu grup yang sama. Yang menjadi satu kelompok terdiri dari sekumpulan data yang semirip mungkin antara data yang satu dengan data yang lain dan harus berbeda dengan data atau objek yang ada pada kelompok lainnya (Jiawei, 2011)

METODE PENELITIAN

Untuk melakukan proses clustering dalam penelitian ini penulis melakukan pengujian dataset berdasarkan metode yang diusulkan dalam penelitian ini yaitu dengan menggunakan metode K-Means Clustering. Adapun dataset yang digunakan adalah data crawling dari TripAdvisor.com, yang pertama kali digunakan oleh Renjith et al (2018) dalam

evaluasi algoritma partisi clustering. Data crawling dapat diakses dari database machine learning pada situs:

Tabel 1. Contoh Travel Reviews Data

User id, atributs
'id1',0.93,1.8,2.29,0.62,0.8,2.42,3.19,2.79,1.82,2.42
'id2',1.02,2.2,2.66,0.64,1.42,3.18,3.21,2.63,1.86,2.32
'id3',1.22,0.8,0.54,0.53,0.24,1.54,3.18,2.8,1.31,2.5
'id4',0.45,1.8,0.29,0.57,0.46,1.52,3.18,2.96,1.57,2.86
'id5',0.51,1.2,1.18,0.57,1.54,2.02,3.18,2.78,1.18,2.54
'id6',0.99,1.28,0.72,0.27,0.74,1.26,3.17,2.89,1.66,3.66
'id7',0.9,1.36,0.26,0.32,0.86,1.58,3.17,2.66,1.22,3.22
'id8',0.74,1.4,0.22,0.41,0.82,1.5,3.17,2.81,1.54,2.88
'id9',1.12,1.76,1.04,0.64,0.82,2.14,3.18,2.79,1.41,2.54
'id10',0.7,1.36,0.22,0.26,1.5,1.54,3.17,2.82,2.24,3.12
'id11',1.47,1.0,7.0,75,1.66,2.76,3.18,2.89,1.66,2.62
'id12',0.96,2.96,0.29,0.38,0.88,2.08,3.17,2.93,1.66,3.42
'id13',0.74,1.44,2.75,0.45,0.98,1.74,3.2,2.87,1.38,2.34
'id14',0.58,1.64,2.27,0.45,1.26,1.72,3.19,2.91,2.3,2.74
'id15',0.96,1.68,2.29,0.51,1.2,2.84,3.2,2.82,2.02,2.46
'id16',1.25,2.52,1.76,0.5,1.46,2.08,3.19,2.74,1.41,2.32
'id17',0.86,1.04,1.76,0.34,0.06,1.1,3.18,2.73,1.15,2.98
'id18',0.61,1.96,2.49,0.66,1.34,1.78,3.2,3.04,1.15,2.42
'id19',0.67,1.36,1.36,0.38,0.82,3.38,3.18,2.86,1.79,2.8
'id20',0.8,1.04,2.1,0.58,1.18,1.98,3.19,2.93,1.22,2.48
'id21',0.86,1.44,1.12,0.41,1.2,2.18,3.18,2.7,1.06,2.94
'id22',0.96,1.44,0.14,0.42,1.06,2.08,3.17,2.74,1.15,3.22
'id23',0.93,1.16,0.29,0.41,1.02,1.36,3.16,2.74,1.34,3.66
'id24',0.54,1.44,0.3,0.41,1.02,1.46,3.17,2.71,1.73,3.04
'id25',1.22,0.96,1.15,0.54,1.1,2.02,3.18,2.77,1.92,2.46
'id26',0.61,2.84,2.8,0.48,0.56,1.52,3.19,2.54,1.6,2.54
'id27',0.77,1.68,2.24,0.75,1.28,2.64,3.19,2.57,1.34,2.38
'id28',0.99,1.28,1.22,0.59,0.94,1.9,3.18,2.79,1.22,2.4
'id29',0.58,1.2,0.18,0.38,0.54,0.76,3.17,2.69,1.63,2.94
'id30',0.64,1.16,3.12,0.45,1.84,3.16,3.2,2.75,1.54,2.46
'id31',0.74,1.28,2.14,0.5,0.94,2.26,3.2,2.78,1.5,2.34
'id32',0.7,2.24,2.32,0.63,0.72,2.12,3.19,2.65,1.28,2.42
'id33',0.64,2.1,6.0,41,2.08,2.22,3.19,2.8,1.76,2.72
'id',1.06,1.12,0.21,0.58,0.98,2.5,3.18,2.9,1.92,2.74
'id35',0.96,1.16,0.45,0.29,0.98,1.42,3.18,2.94,2.02,3.02
'id36',0.88,0.96,1.18,0.43,0.74,1.52,3.18,3.01,1.63,2.94
'id37',0.7,1.28,1.15,0.32,0.4,1.1,3.18,2.78,1.76,2.9
'id38',1.02,1.36,0.91,0.5,0.72,1.22,3.18,2.91,1.92,3.2
'id39',0.7,1.52,2.37,0.62,0.86,1.68,3.2,2.86,1.6,2.58
'id40',1.38,1.08,0.18,0.62,0.14,1.46,3.18,2.88,1.66,2.42
'id41',0.99,1.24,0.15,0.24,1.38,1.98,3.17,2.61,1.02,3.06
'id42',0.58,1.48,0.19,0.41,0.5,0.98,3.16,2.74,1.57,3.38
'id43',0.67,1.24,0.61,0.24,1.06,1.9,3.18,2.82,1.25,2.88
'id44',1.02,2.0,42,0.56,0.46,2.06,3.18,2.86,2.05,2.66
'id45',0.64,1.64,0.34,0.62,0.64,2.22,3.18,3.1,1.98,2.8
'id46',1.7,1.4,0.61,0.39,0.7,1.46,3.17,2.9,1.82,3.38
'id47',0.38,1.08,0.77,0.34,1.06,1.46,3.18,3.15,2.18,3.22
'id48',0.9,1.2,1.24,0.35,0.88,1.86,3.18,2.86,1.47,2.94
'id49',0.45,2.96,0.26,0.4,0.56,1.68,3.18,2.9,1.44,2.72
'id50',0.61,1.08,2.22,0.65,1.2,2.06,3.19,2.62,1.31,2.34
...
...
'id980',0.93,0.56,1.13,0.51,1.34,2.36,3.18,2.87,1.34,2.4

Pada Metode K-Means Clustering, salah satu hal yang cukup penting adalah menentukan jumlah cluster, ini disebabkan jumlah cluster yang berbeda akan menghasilkan nilai euclidean atau jarak data ke setiap titik pusat cluster yang berbeda.

Menentukan Jumlah Atribut.

Atribut dalam penelitian ini ada 11 atribut ditampilkan pada table berikut:

Tabel 2. Atribut Information

Atribut 1	User Id
Atribut 2	Feedback pada galeri seni
Atribut 3	Feedback pada klub dansa
Atribut 4	Feedback pada bar jus
Atribut 5	Feedback pada restoran
Atribut 6	Feedback pada museum
Atribut 7	Feedback pada Resor
Atribut 8	Feedback pada on taman atau tempat pikni
Atribut 9	Feedback pada pantai
Atribut 10	Feedback pada teater
Atribut 11	Feedback pada lembaga keagamaan

Menentukan Nilai Travel Reviews

Setiap traveler memberikan reviews dalam 4 kategori, yaitu: Excellent (4), Very Good (3), Average (2), Poor (1) dan Terrible (0). Data feedback untuk masing-masing rating dapat dilihat dalam tabel berikut:

Tabel 3. Tabel Feedback Travel Reviews

Feedback	Keterangan
4	Excellent
3	Very Good
2	Average
1	Poor
0	Terribel

Travel Reviews Dataset ini terdiri dari sejumlah sample data, dimana jumlah seluruh data terdiri dari 980 data, yang terdiri dari 10 kategori seperti galeri seni, klub dansa, bar jus, restoran, museum, resor , taman atau tempat piknik, pantai, teater, dan lembaga keagamaan dari destinasi wisata di seluruh asia timur.

Inisialisasi Data

Ada beberapa cara yang dapat dilakukan untuk melakukan inisialisasi k pusat cluster, namun yang memberikan nilai awal pada pusat cluster yang paling sering dilakukan, yaitu dengan secara random.

Menentukan Jumlah Cluster

HASIL DAN PEMBAHASAN

Menentukan Cluster Berdasarkan Nilai Euclidean

Peneliti menghitung jarak data ke setiap titik pusat cluster dengan menggunakan teori euclidean yang dirumuskan sebagai berikut:

$$D(i, j) = \sqrt{(X1i - X1j)^2 + (X2i - X2j)^2 + \dots + (Xki - Xkj)^2}$$

Keterangan:

D(i, j) = Jarak data ke i ke pusat cluster j

xki = Data ke i pada atribut data ke k

Xkj = Titik pusat ke j pada atribut ke k

Tabel 4. Hasil Cluster

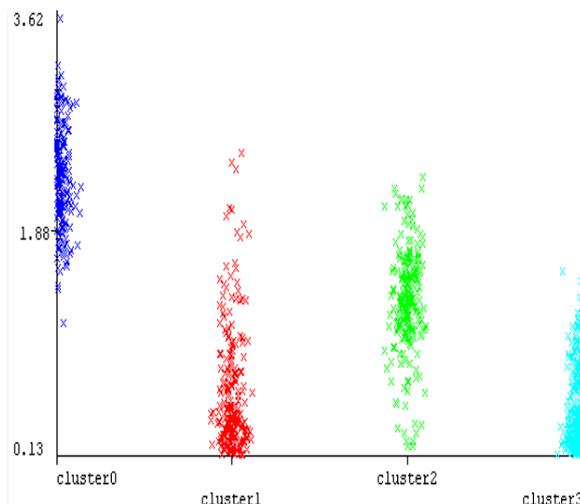
0,0.93,1.8,2.29,0.62,0.8,2.42,3.19,2.79,1.82,2.42,	cluster0
1,1.02,2.2,2.66,0.64,1.42,3.18,3.21,2.63,1.86,2.32,	cluster0
2,1.22,0.8,0.54,0.53,0.24,1.54,3.18,2.8,1.31,2.5,	cluster3
3,0.45,1.8,0.29,0.57,0.46,1.52,3.18,2.96,1.57,2.86,	cluster3
4,0.51,1.2,1.18,0.57,1.54,2.02,3.18,2.78,1.18,2.54,	cluster2
5,0.99,1.28,0.72,0.27,0.74,1.26,3.17,2.89,1.66,3.66,	cluster1
6,0.9,1.36,0.26,0.32,0.86,1.58,3.17,2.66,1.22,3.22,	cluster1
7,0.74,1.4,0.22,0.41,0.82,1.5,3.17,2.81,1.54,2.88,	cluster3
8,1.12,1.76,1.04,0.64,0.82,2.14,3.18,2.79,1.41,2.54,	cluster2
9,0.7,1.36,0.22,0.26,1.5,1.54,3.17,2.82,2.24,3.12,	cluster1
10,1.47,1.0,7.0.75,1.66,2.76,3.18,2.89,1.66,2.62,	cluster2
11,0.96,2.96,0.29,0.38,0.88,2.08,3.17,2.93,1.66,3.42,	cluster1
12,0.74,1.44,2.75,0.45,0.98,1.74,3.2,2.87,1.38,2.34,	cluster0
13,0.58,1.64,2.27,0.45,1.26,1.72,3.19,2.91,2.3,2.74,	cluster0
14,0.96,1.68,2.29,0.51,1.2,2.84,3.2,2.82,2.02,2.46,	cluster0
15,1.25,2.52,1.76,0.5,1.46,2.08,3.19,2.74,1.41,2.32,	cluster0
16,0.86,1.04,1.76,0.34,0.06,1.1,3.18,2.73,1.15,2.98,	cluster1
17,0.61,1.96,2.49,0.66,1.34,1.78,3.2,3.04,1.15,2.42,	cluster0
18,0.67,1.36,1.36,0.38,0.82,3.38,3.18,2.86,1.79,2.8,	cluster2
19,0.8,1.04,2.1,0.58,1.18,1.98,3.19,2.93,1.22,2.48,	cluster0
20,0.86,1.44,1.12,0.41,1.2,2.18,3.18,2.7,1.06,2.94,	cluster2
21,0.96,1.44,0.14,0.42,1.06,2.08,3.17,2.74,1.15,3.22,	cluster1
22,0.93,1.16,0.29,0.41,1.02,1.36,3.16,2.74,1.34,3.66,	cluster1
23,0.54,1.44,0.3,0.41,1.02,1.46,3.17,2.71,1.73,3.04,	cluster1
24,1.22,0.96,1.15,0.54,1.1,2.02,3.18,2.77,1.92,2.46,	cluster2
25,0.61,2.84,2.8,0.48,0.56,1.52,3.19,2.54,1.6,2.54,	cluster0
26,0.77,1.68,2.24,0.75,1.28,2.64,3.19,2.57,1.34,2.38,	cluster0
27,0.99,1.28,1.22,0.59,0.94,1.9,3.18,2.79,1.22,2.4,	cluster2
28,0.58,1.2,0.18,0.38,0.54,0.76,3.17,2.69,1.63,2.94,	cluster1
29,0.64,1.16,3.12,0.45,1.84,3.16,3.2,2.75,1.54,2.46,	cluster0
30,0.74,1.28,2.14,0.5,0.94,2.26,3.2,2.78,1.5,2.34,	cluster0
31,0.7,2.24,2.32,0.63,0.72,2.12,3.19,2.65,1.28,2.42,	cluster0
32,0.64,2.1,6.0.41,2.08,2.22,3.19,2.8,1.76,2.72,	cluster2
33,1.06,1.12,0.21,0.58,0.98,2.5,3.18,2.9,1.92,2.74,	cluster3
34,0.96,1.16,0.45,0.29,0.98,1.42,3.18,2.94,2.02,3.02,	cluster3
35,0.88,0.96,1.18,0.43,0.74,1.52,3.18,3.01,1.63,2.94,	cluster3
36,0.7,1.28,1.15,0.32,0.4,1.1,3.18,2.78,1.76,2.9,	cluster3
37,1.02,1.36,0.91,0.5,0.72,1.22,3.18,2.91,1.92,3.2,	cluster1
38,0.7,1.52,2.37,0.62,0.86,1.68,3.2,2.86,1.6,2.58,	cluster0
39,1.38,1.08,0.18,0.62,0.14,1.46,3.18,2.88,1.66,2.42,	cluster3
40,0.99,1.24,0.15,0.24,1.38,1.98,3.17,2.61,1.02,3.06,	cluster1
41,0.58,1.48,0.19,0.41,0.5,0.98,3.16,2.74,1.57,3.38,	cluster1
42,0.67,1.24,0.61,0.24,1.06,1.9,3.18,2.82,1.25,2.88,	cluster3
43,1.02,2.0,4.2,0.56,0.46,2.06,3.18,2.86,2.05,2.66,	cluster3
44,0.64,1.64,0.34,0.62,0.64,2.22,3.18,3.1,1.98,2.8,	cluster3
45,1.7,1.4,0.61,0.39,0.7,1.46,3.17,2.9,1.82,3.38,	cluster1
46,0.38,1.08,0.77,0.34,1.06,1.46,3.18,3.15,2.18,3.22,	cluster3
47,0.9,1.2,1.24,0.35,0.88,1.86,3.18,2.86,1.47,2.94,	cluster2
48,0.45,2.96,0.26,0.4,0.56,1.68,3.18,2.9,1.44,2.72,	cluster3
49,0.61,1.08,2.22,0.65,1.2,2.06,3.19,2.62,1.31,2.34,	cluster0
50,1.02,1.2,0.3,0.38,1.1,1.9,3.17,2.93,1.34,3.42,	cluster1
....	
....	
979,0.93,0.56,1.13,0.51,1.34,2.36,3.18,2.87,1.34,2.4,	cluster2

Final Cluster Centroid

Proses clustering berhenti apabila pusat cluster tidak berubah. Berikut hasil cluster centroid.

Tabel 5. Final cluster Centroids

Atribut	Cluster 0	Cluster 1	Cluster 2	Cluster 3
Atribut 1	0.8602	0.9487	0.9576	0.8258
Atribut 2	1.4892	1.3118	1.34	1.3178
Atribut 3	2.2722	0.5935	1.3232	0.431
Atribut 4	0.6005	0.3925	0.5724	0.5712
Atribut 5	1.1572	0.7464	1.2011	0.7847
Atribut 6	2.1746	1.4423	2.2119	1.7044
Atribut 7	3.1932	3.1728	3.1818	3.1797
Atribut 8	2.782	2.8258	2.829	2.8745
Atribut 9	1.5869	1.4526	1.4849	1.7041
Atribut 10	2.4719	3.2299	2.6648	2.7516



Gambar 1. Data setelah dilakukan cluster

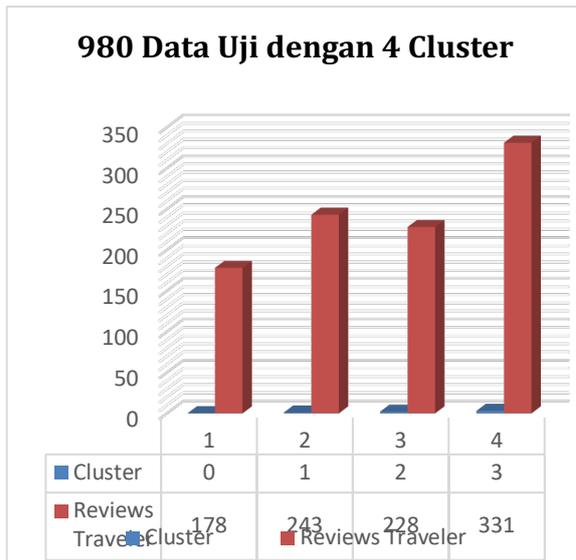
Cluster Instance

Cluster instances ditunjukkan pada Tabel 6 berikut:

Tabel 6. Clustered Instances

Cluster	Percentage
0	178 (18%)
1	243 (25%)
2	228 (23%)
3	331 (34%)

Grafik perbandingan hasil setiap cluster dari 980 Data uji dengan 4 Cluster ditampilkan pada Gambar 2 berikut:



Gambar 2. Grafik hasil cluster dari 980 data

KESIMPULAN

Dari penelitian yang telah dilakukan, maka dapat ditarik kesimpulan sebagai berikut:

1. Travel reviews dataset yang terdiri dari 980 reviews traveler, dengan 10 atribut (kategori destinasi wisata), diperoleh 4 cluster ($k=4$)
2. Cluster pertama terdiri dari 178 traveler reviews atau 18%, yaitu feedback pada galeri seni = 0.8602, feedback pada club dansa = 1.4892, feedback bar jus = 2.2722, feedback pada restoran = 0.6005, feedback pada museum = 1.1572, feedback on resor = 2.1746, feedback pada taman atau tempat piknik = 3.1932, feedback pada pantai = 2.782, feedback pada teater = 1.5869, feedback pada lembaga keagamaan = 2.4719
3. Cluster kedua terdiri dari 243 traveler reviews, atau 25%, yaitu feedback pada galeri seni = 0.9487, feedback pada club dansa = 1.3118, feedback pada bar jus = 0.5935, feedback pada restoran = 0.3925, feedback on museum = 0.7464, feedback pada resor = 1.4423, feedback pada taman atau tempat piknik = 3.1728, feedback pada pantai = 2.8258, feedback pada teater = 1.4526, feedback pada lembaga keagamaan = 3.2299
4. Cluster ketiga terdiri dari 228 traveler reviews, atau 23%, yaitu feedback pada galeri seni = 0.9576, feedback pada club dansa = 1.34, feedback pada bar jus = 1.3232, feedback pada restoran = 0.5724, feedback pada museum = 1.2011, feedback on resor = 2.2119, feedback pada taman atau tempat piknik = 3.1818, feedback pada pantai = 2.829, feedback pada teater =

1.4849, feedback pada lembaga keagamaan = 2.6648

5. Cluster keempat terdiri dari 331 traveler reviews, atau 34%, yaitu feedback pada galeri seni = 0.8258, feedback pada club dansa = 1.3178, feedback pada bar jus = 0.431, feedback pada restoran = 0.5712, feedback pada museum = 0.7847, feedback pada resor = 1.7044, feedback pada taman atau tempat piknik = 3.1797, feedback pada pantai = 2.8745, feedback pada teater = 1.7041, feedback pada lembaga keagamaan = 2.7516

DAFTAR PUSTAKA

- Agusta, Y. (2007). K-Means Penerapan, Permasalahan dan Metode Terkait. *Jurnal Sistem dan Informatika*, 3(1), 47-60.
- Almana, A. M. dan Abdulrahman A. M. (2013). The Impact of Electronic Word of Mouth on Consumers Purchasing Decisions. *International Journal of Computer Applications*, 82, 34-41.
- Berry, M. W and Browne, M. (2006). Lecture notes in data mining. *World Scientific*.
- Chen, Y. dan Jinhong X. (2008). Online Consumer Ulasan: Word of Mouth sebagai Baru Unsur Bauran Pemasaran Komunikasi. *Manajemen Sains*, 54(3), 477-489.
- Jiawei, H. (2011). *Data Mining: Concepts and Techniques*. 3rd Edition. Amsterdam: Morgan Kaufmann.
- Llyod, S. (1982). Least Square Quantization in PCM. *IEEE Transactions on Information Theory*, 28(2), 129-137.
- Lee, E. J. and Shin, S. Y. (2014). When do Consumer buy online product reviews? Effects of review quality, product type, and reviewer's photo. *Computers in Human Behavior*, 31, 356-366
- Reddy and Jana (2012). Initialization K-means clustering using voronoi diagram. *Procedia Technology*, 4, 395-400.
- Susanto, S. dan Suryadi, D. (2010). *Pengantar data mining: menggali pengetahuan dari bongkahan data*. Yogyakarta: Penerbit Andi.