JURNAL

ANALISA SEBARAN WILAYAH PENDENGAR ACARA HELLO DANGDUT RADIO WIJANGSONGKO KEDIRI DENGAN METODE K-MEANS

AREA ANALYSIS OF THE LISTENERS HELLO DANGDUT SHOW RADIO WIJANG SONGKO KEDIRI WITH K-MEANS METHOD

Oleh:

NUHDI FUTUHAL ARIFIN

NPM. 12.1.03.02.0443

Dibimbing oleh :

1. **Ahmad Bagus Setiawan, ST, M.Kom, MM**
2. **Fajar Rohman Hariri, M.Kom**

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS TEKNIK

UNIVERSITAS NUSANTARA PGRI KEDIRI

2017

SURAT PERNYATAAN
ARTIKEL SKRIPSI TAHUN 2017

Yang bertanda tangan di bawah ini:

Nama Lengkap	: Nuhdi Futuhal Arifin
NPM	: 12.1.03.02.0443
Telepun/HP	: 081231895897
Alamat Surel (Email)	: mbahbejorws@gmail.com
Judul Artikel	: Analisa Sebaran Wilayah Pendengar Acara Hello Dangdut Radio Wijang Songko Kediri Dengan Metode K-Means
Fakultas – Program Studi	: Fakultas Teknik – Prodi Teknik Informatika
Nama Perguruan Tinggi	: Universitas Nusantara PGRI Kediri
Alamat Perguruan Tinggi	: Mojoroto Gang I No. 6 Kota Kediri

Dengan ini menyatakan bahwa :

a. Artikel yang saya tulis merupakan karya saya pribadi (bersama tim penulis**)** dan bebas plagiarisme;
b. Artikel telah diteliti dan disetujui untuk diterbitkan oleh Dosen Pembimbing I dan II.

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

Mengetahui		Kediri, 30-01-2017
Pembimbing I Ahmad Bagus Setiawan,ST,M.Kom,MM NIDN. 0703018704	Pembimbing II Fajar Rohman Hariri,M.Kom NIDN. 0715058902	Penulis, Nuhdi Futuhal Arifin NPM:12.1.03.02.0443

Skripsi oleh :

NUHDI FUTUHAL ARIFIN

NPM. 12.1.03.02.0443

Judul:

ANALISA SEBARAN WILAYAH PENDENGAR ACARA HELLO DANGDUT RADIO WIJANG SONGKO KEDIRI DENGAN METODE K-MEANS

Telah dipertahankan di depan Panitia Ujian/ Sidang Skripsi

Program Studi Teknik Informatika Fakultas Teknik

Universitas Nusantara PGRI Kediri

Pada tanggal: 17 Januari 2017

Dan Dinyatakan Telah Memenuhi Persyaratan

PanitiaPenguji :

1. Ketua : Ahmad Bagus Setiawan, S.T., M.M., M.Kom.
2. Penguji I : Ardi Sanjaya, M.Kom.
3. Penguji II : Fajar Rohman Hariri, M.Kom

Mengetahui,
Dekan Fakultas Teknik

Dr. Suryo Widodo, M.Pd
NIP. 19640202 199103 1002

ANALISA SEBARAN WILAYAH PENDENGAR ACARA HELLO DANGDUT RADIO WIJANG SONGKO KEDIRI DENGAN METODE K-MEANS

Nuhdi Futuhal Arifin
NPM: 12.1.03.02.0443
Fakultas Teknik – Prodi Teknik Informatika
Email : mbahbejorws@gmail.com
Dosen Pembimbing 1: Ahmad Bagus Setiawan, ST, M.Kom, MM
Dosen Pembimbing 2 : Fajar Rohman Hariri, M/Kom
UNIVERSITAS NUSANTARA PGRI KEDIRI

Abstrak

Nuhdi Futuhal Arifin : Analisa Sebaran Wilayah Pendengar Acara Hello Dangdut Radio Wijang Songko Kediri Dengan Metode K-Means, Skripsi, Teknik Informatika Fakultas Teknik Universitas Nusantara PGRI Kediri

Penelitian ini dilatarbelakangi hasil pengamatan dan pengalaman peneliti, bahwa acara Hello Dangdut (Helda) merupakan acara unggulan Radio Wijang Songko dengan menyajikan content full humor yang sangat digemari oleh pendengar, sehingga banyak pemasang iklan ingin spot produknya tayang di acara ini.

Permasalahan penelitian ini adalah (1) Di wilayah manakah sebaran pendengar dengan kategori SMS tinggi, sedang dan rendah di acara Hello Dangdut Radio Wijang Songko pada bulan Juli dan Agustus 2016 ?

Penelitian ini menggunakan metode K-Means yang digunakan untuk mempartisi data ke dalam kelompok tertentu sehingga data yang paling dekat kesamaannya akan dikelompokkan ke dalam satu cluster/kelompok yang sama, sedangkan data dengan karakteristik berbeda akan dikelompokkan dalam kelompok tersendiri.

Kesimpulan hasil penelitian ini adalah (1) Metode K-Means dapat digunakan dalam mengelompokkan sebaran data SMS berdasarkan kategori wilayah pendengar acara Hello Dangdut Radio Wijang Songko Kediri. (2) Dari hasil clustering dapat diekstrak informasi pada tiap kota bisa diketahui wilayah yang termasuk kategori pengisim SMS tinggi, sedang dan rendah. (3) Pengujian hasil cluster dengan Silhuette Coefficient menunjukkan bahwa hasil clustering menggunakan metode K-Means sudah terkelompok dengan tepat.

Berdasarkan simpulan hasil penelitian ini direkomendasikan: (1) Dalam teknik visualisasi data bisa dilakukan dalam bentuk integrasi peta wilayah, sehingga user bisa mendapatkan gambaran visualisasi data yang lebih mudah. (2) Dapat dilakukan pengembangan program dengan data yang lebih banyak dalam rentang bulan maupun wilayah yang diteliti.

Kata Kunci : clustering K-Means, sebaran wilayah, Silhouette Coefficient, Data Mining

I. LATAR BELAKANG

Acara Hello Dangdut (Helda) merupakan acara unggulan Radio Wijang Songko. Banyak pemasang iklan ingin spot produknya tayang di acara ini. Pertanyaan yang sering dilontarkan oleh pemasang iklan adalah dari wilayah mana saja pendengar acara Hello Dangdut berasal.

Hal ini merupakan pertanyaan utama karena pemasang iklan pasti ingin spot produknya diputar pada acara dengan cakupan pendengar yang tepat sehingga memberikan imbas yang positif untuk penjualan produknya.

Dari latar belakang permasalahan tersebut penulis membuat penelitian mengenai Analisa Sebaran Wilayah Pendengar Acara Hello Dangdut Radio Wijang Songko Kediri

II. METODE

Dalam masalah ini peneliti menggunakan algoritma K-Means untuk Mengelompokkan objek dengan membagi data ke dalam cluster-cluster yang ada sejumlah k partisi. Dimana k merupakan banyaknya cluster yang berasal dari input data.

Untuk menguji bahwa hasil cluster menggunakan K-Means sudah tepat digunakan metode *silhouette coefficient* untuk melihat kualitas dan kekuatan cluster serta seberapa baik suatu objek ditempatkan dalam suatu cluster.

K-Means termasuk dalam *partitioning clustering* yaitu suatu metode penganalisaan data atau metode data mining yang melakukan proses permodelan tanpa supervisi (Nasari dkk, 2015). K-Means memisahkan data ke *k* daerah bagian yang terpisah, dimana *k* adalah bilangan integer positif. Algoritma K-Means sangat terkenal karena kemudahan dan kemampuannya untuk mengklasifikasi data besar dan *outlier* dengan sangat cepat.

Berikut adalah langkah-langkah algoritma *K-Means* :

1. Penentuan pusat cluster awal

   Dalam menentukan n buah pusat *cluster* awal dilakukan pembangkitan bilangan random yang merepresentasikan urutan data *input*. Pusat awal *cluster* didapatkan dari data sendiri bukan dengan menentukan titik baru, yaitu dengan merandom pusat awal dari data.
2. Perhitungan jarak dengan pusat cluster

   Untuk mengukur jarak antara data dengan pusat *cluster* digunakan *Euclidian distance*.

   Algoritma perhitungan jarak data dengan pusat *cluster*

a) Ambil nilai data dan nilai pusat *cluster*

b) Hitung *Euclidian distance* data dengan tiap pusat *cluster* dengan rumus

$$d_{ij} = \sqrt{\sum_{k=1}^{p} \{x_{ik} - x_{jk}\}^2}$$

d_{ij} = Jarak objek antara objek i dan j

P = Dimensi data

X_{ik} = Koordinat dari obyek i pada dimensi k

X_{jk} = Koordinat dari obyek j pada dimensi k

3. Pengelompokan data

Jarak hasil perhitungan akan dilakukan perbandingan dan dipilih jarak terdekat antara data dengan pusat *cluster*, jarak ini menunjukkan bahwa data tersebut berada dalam satu kelompok dengan pusat *cluster* terdekat.

Algoritma pengelompokkan data

a) Ambil nilai jarak tiap pusat *cluster* dengan data

b) Cari nilai jarak terkecil

c) Kelompokkan data dengan pusat *cluster* yang memiliki jarak terkecil.

4. Penentuan pusat cluster baru

Untuk mendapatkan pusat *cluster* baru bisa dihitung dari rata-rata nilai anggota *cluster* dan pusat *cluster*. Pusat *cluster* yang baru digunakan untuk melakukan iterasi selanjutnya, jika hasil yang didapatkan belum konvergen. Proses iterasi akan berhenti jika telah memenuhi maksimum iterasi yang dimasukkan oleh *User* atau hasil yang dicapai sudah konvergen (pusat *cluster* baru sama dengan pusat *cluster* lama).

Algoritma penentuan pusat *cluster* baru

a) Cari jumlah anggota tiap *cluster*

b) Hitung pusat baru dengan mencari nilai rata-rata data dari anggota cluster.

Setelah cluster terbentuk untuk menguji keakuratan cluster digunakan *silhouette coefficient* :

$$Si = (b_i - a_i) / \max(a_i, b_i)$$

Nilai *silhouette coefficient* dapat bervariasi antara -1 hingga 1. Hasil *clustering* dikatakan baik jika nilai *silhouette coefficient* bernilai positif ($a_i < b_i$) dan a_i mendekati 0, sehingga akan menghasilkan nilai *silhouette coefficient* yang maksimum yaitu 1 saat $a_i = 0$. Dengan demikian, jika s_i=1 berarti objek *i* sudah berada dalam *cluster* yang tepat. Jika nilai s_i= 0 maka objek *i* berada diantara dua *cluster* sehingga objek tersebut tidak jelas harus dimasukkan ke dalam *cluster* A atau *cluster* B. Akan

tetapi, jika s_i= -1 artinya struktur *cluster* yang dihasilkan *overlapping*, sehingga objek *i* lebih tepat dimasukkan ke dalam *cluster* yang lain. Nilai rata-rata *silhouette* dari tiap objek dalam suatu *cluster* dalah suatu ukuran yang menunjukkan seberapa ketat data dikelompokkan dalam *cluster* tersebut. Berikut ini ukuran nilai *silhouette* berdasarkan Kaufman dan Rousseeuw.

Nilai *silhouette coefficient* (SC)

0.7 < SC <= 1 strong structure

0.5 < SC <= 0.7 medium structure

0.25 < SC <= 0.5 weak structure

SC <= 0.25 no structure

Studi Kasus Analisa Sebaran Wilayah Dengan K-Means

No	Wilayah	JUMLAH SMS
1	Semampir	22
2	Tawangsari	43
3	Ngadirejo	19
4	Pakelan	21
5	Sananwetan	17
6	Banjaran	17
7	Rejoagung	18
8	Kemasan	25
9	Gandu	19
10	Gedog	18
11	Bulurejo	22
12	Pocanan	26
13	Kayangan	24
14	Klampok	22
15	Jagalan	26
16	Cukir	19
17	bendogerit	21
18	kaliombo	18
19	plosokerep	20
20	Jatirejo	27
21	ngronggo	12
22	ringinpitu	12
23	jatipelem	11
24	bandung	9
25	bagorkulon	11
26	dandangan	11
27	Diwek	10
28	kedungwaru	16
29	karangtengah	13
30	banaranwetan	11
31	Ceweng	9
32	gemenggeng	12
33	karang tengah	10
34	Ketanon	12
35	Buduran	10

Untuk pengelompokan data wilayah SMS pendengar menggunakan algoritma k-means, dibagi dalam 3 (tiga) kelompok, dengan centroid awal yang ditentukan secara acak adalah C1 = 33, C2 = 21, C3 = 9

Dari hasil penghitungan iterasi 1 didapatkan kelompok wilayah sebagai berikut :

C1	C2	C3
tawangsari	semampir	ngronggo
jatirejo	ngadirejo	ringinpitu
	pakelan	jatipelem
	sananwetan	bandung
	banjaran	bagorkulon
	rejoagung	dandangan
	kemasan	diwek
	gandu	karangtengah

	Gedog	banaranwetan
	Bulurejo	ceweng
	Pocanan	gemenggeng
	kayangan	karang tengah
	Klampok	ketanon
	Jagalan	buduran
	Cukir	
	bendogerit	
	kaliombo	
	plosokerep	
	kedungwaru	

Perhitungan pusat klaster baru dilakukan dengan mencari nilai rata-rata data nilai wilayah SMS pendengar dari sebanyak anggota pada kelompok C, sehingga diperoleh pusat klaster baru C1 = 35 C2 = 20.53 C3 = 10.93.

Dari hasil penghitungan iterasi 2 didapatkan kelompok wilayah sebagai berikut :

C1	C2	C3
tawangsari	semampir	ngronggo
	ngadirejo	ringinpitu
	Pakelan	jatipelem
	sananwetan	bandung
	Banjaran	bagorkulon
	rejoagung	dandangan
	Kemasan	diwek
	Gandu	karangtengah
	Gedog	banaranwetan
	Bulurejo	ceweng
	Pocanan	gemenggeng
	kayangan	karang Tengah
	Klampok	ketanon
	Jagalan	buduran
	Cukir	
	bendogerit	
	kaliombo	
	plosokerep	
	Jatirejo	
	kedungwaru	

Karena ada data yang berpindah cluster maka dilakukan penentuan pusat cluster baru lagi dengan mencari nilai rata-rata data nilai wilayah SMS pendengar dari sebanyak anggota pada kelompok C, sehingga diperoleh pusat klaster baru C1 = 43 C2 = 20.85 C3 = 10.93.

Dari hasil penghitungan iterasi 3 didapatkan kelompok wilayah sebagai berikut :

C1	C2	C3
tawangsari	semampir	ngronggo
	ngadirejo	ringinpitu
	pakelan	jatipelem
	sananwetan	bandung
	banjaran	bagorkulon
	rejoagung	dandangan
	kemasan	diwek
	gandu	karangtengah
	gedog	banaranwetan
	bulurejo	ceweng
	pocanan	gemenggeng
	kayangan	karang Tengah
	klampok	ketanon
	jagalan	buduran
	cukir	
	bendogerit	
	kaliombo	
	plosokerep	
	jatirejo	
	kedungwaru	

Karena posisi anggota cluster tidak mengalami perubahan lagi dengan hasil yang sama dengan iterasi 2, maka proses cluster berhenti. Sehingga hasil akhir dari pengklasteran sebaran wilayah pendengar ini adalah hasil cluster dari iterasi 3.

Pengujian Silhuette Coefficient

kelompok C1

No	Kelompok	Nilai SC	Kategori
1	Tawangsari	1	Strong

Pada pengujian cluster C1 menghasilkan nilai Strong yakni “1” karena nilai pada cluster tersebut sama persis dengan titik centroid terakhir.

Pengujian Silhuette Coefficient kelompok C2

No	Kelompok	Nilai SC	Kategori
1	semampir	0.437328454	Weak
2	ngadirejo	0.528819762	Medium
3	pakelan	0.528819762	Medium
4	sananwetan	0.62031107	Medium
5	banjaran	0.711802379	Strong
6	rejoagung	0.894784995	Strong
7	kemasan	0.894784995	Strong
8	Gandu	0.894784995	Strong
9	Gedog	0.986276304	Strong
10	bulurejo	0.986276304	Strong
11	pocanan	0.922232388	Strong
12	kayangan	0.83074108	Strong
13	klampok	0.83074108	Strong
14	jagalan	0.83074108	Strong
15	Cukir	0.739249771	Strong
16	bendogerit	0.739249771	Strong
17	kaliombo	0.739249771	Strong
18	plosokerep	0.647758463	Strong
19	jatirejo	0.647758463	Strong
20	kedungwaru	0.556267155	Strong

Hasil pengujian cluster C2 didapatkan 16 anggota cluster dengan nilai "Strong", 3 anggota cluster dengan nilai "Medium" dan satu angota kelompok dengan nilai "Weak"

Pengujian Silhuette Coefficient kelompok C3

No	Kelompok	Nilai SC	Kategori
1	ngronggo	0.948681	Strong
2	ringinpitu	0.900719	Strong
3	jatipelem	0.948681	Strong
4	bandung	0.948681	Strong
5	bagorkulon	0.948681	Strong
6	dandangan	0.996643	Strong
7	Diwek	0.996643	Strong
8	Karangtengah	0.996643	Strong
9	Banaranwetan	0.996643	Strong
10	Ceweng	0.955396	Strong
11	Gemenggeng	0.955396	Strong
12	karang tengah	0.955396	Strong
13	Ketanon	0.907434	Strong
14	Buduran	0.907434	Strong

Pada pengujian cluster C3 didapatkan hasil cluster yang bagus yakni bernilai "Strong" untuk semua anggota clusternya.

Kesimpulan Pengujian

Berdasarkan hasil pengujian dengan Silhuette Coefficient bisa diambil kesimpulan bahwa hasil clustering K-Means sudah terkelompokkan dengan tepat, hal ini dibuktikan dengan penghitungan Silhuette Coefficient dengan hasil akhir nilai S (Strong) pada tiap kelompok lebih dominan dari pada M (Medium) dan L (lower key)

Analisa Data :

Wilayah pendengar dibagi dalam 5 Kota, yang masing – masing kota mempunyai kelompok wilayah tersendiri, seperti pada tabel berikut :

No	Kota	Kelompok Wilayah
1.	Kediri	semampir, dandangan, ngadirejo, pakelan, pocanan, banjaran, jagalan, kemasan, kaliombo, ngronggo
2.	Nganjuk	bagorkulon, balongrejo, banarankulon, banaranwetan, buduran, gandu, gemenggeng, girirejo, karang tengah
3.	Jombang	balongbesuk, bandung, bendet, brambang, bulurejo, ceweng, cukir, diwek, jatipelem, jatirejo, kayangan
4.	Tulungagung	tunggulsari,

		ringinpitu, kedungwaru, rejoagung, ketanon, tawangsari, mangunsari, bangoan
5.	Blitar	bendogerit, plosokerep, klampok, gedog, karangtengah, sananwetan

Maka data hasil clustering diatas bisa diekstrak informasi sebagai berikut :

1. Wilayah Kediri
   a. Kelompok interaksi SMS sedang (Semampir, ngadirejo, pakelan, banjaran, kemasan, pocanan, jagalan, kaliombo)
   b. Kelompok interaksi SMS rendah (Ngronggo, dandangan)
2. Wilayah Nganjuk
   a. Kelompok interaksi SMS sedang (Gandu)
   b. Kelompok interaksi SMS rendah (Bagor kulon, karang tengah, banaran wetan, gemenggeng, buduran)
3. Wilayah Jombang
   a. Kelompok interaksi SMS sedang (Bulurejo, kayangan, cukir)
   b. Kelompok interaksi SMS rendah (Jatipelem, bandung, diwek, ceweng)
4. Wilayah Tulungagung
   a. Kelompok interaksi SMS Tinggi (Tawangsari)
   b. Kelompok interaksi SMS sedang (Rejoagung, kedungwaru)
   c. Kelompok interaksi SMS rendah (Ringin pitu, Ketanon)
5. Wilayah Blitar
   Kelompok interaksi SMS kategori sedang (Sananwetan, gedog, klampok, bendogerit, plosokerep)

III. HASIL DAN KESIMPULAN

Dari hasil pembahasan pada bab sebelumnya maka dapat disimpulkan :

1. Metode K-Means dapat digunakan dalam mengelompokkan sebaran data SMS berdasarkan kategori wilayah pendengar acara Hello Dangdut Radio Wijang Songko Kediri
2. Dari hasil clustering dapat diekstrak informasi pada tiap kota bisa diketahui wilayah yang termasuk kategori pengisim SMS tinggi, sedang dan rendah
3. Pengujian hasil cluster dengan Silhuette Coefficient menunjukkan bahwa hasil clustering menggunakan metode K-Means sudah terkelompok dengan tepat
4. Keakuratan sebaran wilayah pendengar sangat tergantung pada keakuratan pengumpulan data SMS pendengar acara Hello Dangdut Radio Wijang Songko Kediri

IV. DAFTAR PUSTAKA

1. Agusta, Yudi, PhD. 2007. *K-Means Penerapan, Permasalahan dan Metode Terkait*. Jurnal Sistem Informatika Vol.3, 47-60.
2. Aspriyono, Hari, 2011, *Tutorial Singkat Aplikasi Database Berbasis Web.* Bengkulu : Tutorial Gratis.
3. Ediyanto, Muhlasah Novitasari Mara, Neva Satyahadewi. 2013. *Pengklasifikasian Karakteristik Dengan Metode K-Means Cluster Analysis*. Bimaster Volume 2, No. 2, 133 – 136.
4. Handoyo, Reny, R. Rumani M, Surya Michrandi Nasution, 2014. *Perbandingan Metode Clustering Menggunakan Metode Single Linkage dan K-Means Pada Pengelompokan Data.* Bandung : JSM STMIK Mikroskil.
5. Ilmu Biner, *Belajar Mudah Algoritma Data Mining Clustering K=Means.* (Online). Tersedia : http://www.ilmubiner.blogspot.com, diunduh 12 Agustus 2016
6. Kusrini, dan Emha Taufik Luthfi. 2009. *Algoritma Data Mining*. Yogya,karta : Penerbit Andi.
7. Nasari, Fina, Surya Darma. 2015. *Penerapan K-Means Clustering Pada data Penerimaan Mahasiswa Baru (Studi Kasus : Universitas Potensi Utama).* Yogyakarta : STMIK AMIKOM.
8. Nugroho, Bunafit, 2014. *Aplikasi Sistem Pakar Dengan PHP dan Editor Dreamweaver*. Yogyakarta : Penerbit Gava Media
9. Raharjo, Budi, 2015. *Belajar Otodidak MySql (Teknik Pembuatan dan Pengelolaan Database)*. Bandung : Penerbit Informatika.
10. Rohmawati W, Nurul, Sofi Defiyanti, Mohamad Jajuli, 2015. *Implementasi Algoritma K-Means Dalam Pengklasteran Mahasiswa Pelamar Beasiswa.* Karawang : Jurnal Ilmiah Teknologi Informasi Terapan, Volume I.
11. Simamora B. 2005. *Analisis Multivariat Pemasaran.* Jakarta : PT. Gramedia Pustaka Utama.
12. Sumadyo, Malikus, *Koefisien Silhouette, Metode Pengukuran Efektifitas Algoritma Klastering*. Bekasi : Teknik Komputer Universitas Islam '45'.