

PENGUNAAN INTERNET DIKALANGAN SISWA SD DI KOTA TERNATE: SUATU *SURVEY*, PENERAPAN ALGORITMA *CLUSTERING* DAN VALIDASI DBI

Firman Tempola^{*1}, Miftah Muhammad², Abdul Mubarak³

^{1,3}Teknik Informatika Universitas Khairun Ternate, ²Teknik Elektro Universitas Khairun Ternate
Email: ¹firman.tempola@unkhair.ac.id, ²miftahmuh@unkhair.ac.id, ³amuba@unkhair.ac.id

^{*}Penulis Korespondensi

(Naskah masuk: 14 Agustus 2019, diterima untuk diterbitkan: 26 November 2020)

Abstrak

Penggunaan internet dimasyarakat global terus tumbuh, tak hanya terjadi pada masyarakat dewasa melainkan juga pada anak-anak. Internet tidak hanya berdampak pada hal positif melainkan juga pada hal negatif. Di Ternate penggunaan internet terus tumbuh hal ini karena semakin mudah dalam mengakses internet. Namun laporan secara ilmiah mengenai penggunaan internet di Kota Ternate belum ada. Untuk itu, bagaimana mengetahui penggunaan internet dikalangan anak SD di kota Ternate. Penelitian itu bertujuan untuk mencari tahu penggunaan internet di Kota Ternate dengan cara survey secara langsung kepada kalangan anak SD di kota Ternate. Selain itu, data-data dari hasil survey kemudian di *cluster* dengan menggunakan algoritma *k-means clustering*. kemudian dilakukan validasi *clustering* dengan *davies bouldin index*. Hasil dari penelitian ini dari 933 responden diperoleh 51,45 % siswa SD di kota Ternate aktif di jejaring sosial dengan 53,70% di *whatsapp*, 40,30% di *instagram* dan 27,80% di *facebook*. Untuk aktivitas ketika membuka youtube terdapat 61,60% sering menonton video di youtube dengan 61,60% video karton, komedi 49,80% dan konten edukasi 28,40%. Sedangkan untuk game online, yang aktif dalam bermain game online yaitu 49,41%. Untuk penerapan algoritma *clustering k-means* pada 32 sekolah SD di Kota Ternate diperoleh cluster terbaik saat pembagian 4 cluster, hal ini berdasarkan nilai *davies bouldin index* yang diperoleh sebesar 0,773 lebih kecil dibandingkan dengan pembagian cluster lainnya.

Kata kunci: *Pengguna internet, Siswa SD, Clustering K-Means, Validasi DBI*

USE OF THE INTERNET IN THE ELEMENTARY SCHOOL STUDENTS IN TERNATE CITY: A SURVEY, IMPLEMENTED OF CLUSTERING ALGORITHM AND VALIDATION DBI

Abstract

The use of the internet in the global community continues to grow, not only in adults but also in children. The internet does not only have positive effects but also negative things. In Ternate the use of the internet continues to grow because it is easier to access the internet. However, scientific reports regarding the use of the internet in the city of Ternate do not yet exist. for that, how to find out the use of the internet among elementary school children in the city of Ternate. The research aims to find out the use of the internet in the city of Ternate by means of a direct survey among elementary school children in the city of Ternate. In addition, the data from the survey results are then clustered using the *k-means clustering algorithm*. Then the clustering validation was performed with the *bouldin index davies*. The results of this study of 933 respondents obtained 51.45% of elementary school students in Ternate were active in social networks with 53.70% on *whatsapp*, 40.30% on *Instagram* and 27.80% on *Facebook*. For activities when opening *YouTube* there are 61.60% often watching videos on *YouTube* with 61.60% cardboard videos, comedy 49.80% and educational content 28.40%. As for online games, those active in playing online games are 49.41%. For the application of the *k-means clustering algorithm* in 32 elementary schools in Ternate, the best cluster was obtained when the division of 4 clusters, this was based on the *bouldin index davies* value obtained by 0.773 smaller than the other cluster divisions.

Keywords: *Internet users, elementary students, clustering K-Means, validation DBI*

1. PENDAHULUAN

Di era globalisasi saat ini, Internet menjadi salah satu hal yang sangat dibutuhkan. Hampir

sebagian besar aktivitas dalam kehidupan sehari-hari selalu berhubungan dengan yang namanya internet. Penggunaan internet juga tidak hanya

terbatas pada masyarakat dewasa melainkan juga pada kalangan anak-anak. Hal ini karena sangat mudah dan terjangkau dalam mengakses internet.

Pertumbuhan penggunaan internet di Indonesia terus mengalami peningkatan, sebagaimana yang dilaporkan oleh (Internetstats, 2019) bahwa pada tahun 2000 penggunaan internet di Indonesia hanya 2.000.000 pengguna, namun laporan terakhir pada tahun 2019 meningkat sangat signifikan yaitu berjumlah 143.260.000 dengan jumlah penduduk 266.794.980 jiwa atau sekitar 53,69 % dari jumlah penduduk Indonesia sebagai pengguna internet. Selain itu laporan dari European Union Kids Online dan dipublikasikan oleh European Commission's Safer internet Programme menyebutkan bahwa penggunaan internet di kalangan anak-anak 83% bermain game dan 71% aktif di jejaring social (Internetsociety, 2017)

Penelitian terkait pengguna internet dikalangan anak-anak Di Indonesia pernah dilakukan oleh (Chandra, 2013)) menyebutkan 12% anak-anak di Surabaya usia 5 tahun sudah mengenal internet. Berbeda dengan (Chandra, 2013), (Husni and Fatulloh, 2016) melakukan kategorisasi pengguna internet dikalangan SD dan SMP di kota Bandung dimana 31% akses internet ketika berada di rumah dengan akses melalui perangkat Handphone 34,32%. Hasil survey lain juga menunjukkan 26,71% penggunaan internet untuk mencari tugas, namun perbedaan dengan sering mengakses jejaring sosial dan bermain game tidak signifikan yaitu 22,51% (Jejaring sosial) dan 22,99% (Game). Namun dari kedua penelitian tersebut tidak ada tahapan uji reliabilitas dan validitas pertanyaan yang sebutkan dalam artikelnya.

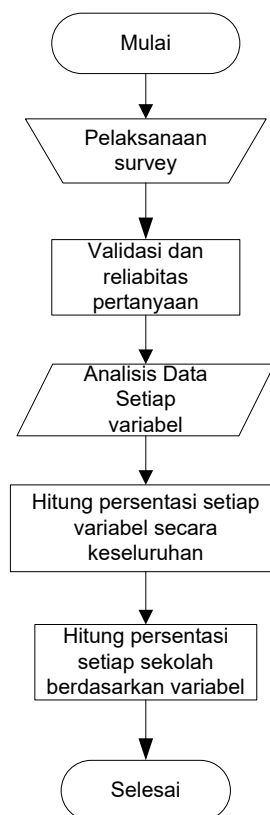
Di Ternate penggunaan internet dikalangan anak-anak khususnya pada anak-anak di usia SD juga terus meningkat. Salah satu faktornya adalah mudah dalam mengakses internet. terlihat dari *Indihome* yang merambah ke rumah-rumah warga hal ini berdasarkan laporan dari media lokal malutpost pada tanggal 10 Juli 2018, untuk itu penting sebuah survey agar dapat memberi rekomendasi kepada para orang tua agar mengontrol anak-anak dirumah agar tidak salah dalam mengakses internet untuk hal negatif mengingat 86 % anak-anak dalam mengakses internet ketika berada di rumah (Nces, 2018). Pada penelitian ini juga selain melakukan survey penggunaan internet di kalangan siswa SD di kota Ternate juga melakukan penerapan algoritma clustering sekolah SD di kota ternate dengan K-Means.

Penerapan algoritma *clustering* k-means telah di terapkan diberbagai kasus-kasus saat ini sebagaimana yang dilakukan oleh (Aprianti and Permadi, 2018) pada data kecelakaan, (Romdhoni dkk, 2018) pada mobilitas moda transportasi, dan (Burhan, Mahmudy and Dermawi, 2018) pada prediksi curah hujan. Setelah penerapan algoritma *clustering*, perlu diketahui juga kualitas dari *cluster*

yang terbentuk. Untuk itu, dalam mengukur kualitas *cluster* pada penelitian ini menggunakan *Davies Bouldin Index* (DBI). DBI pernah dimanfaatkan oleh (Jumadi Dehotman Sitompul, Salim Sitompul and Sihombing, 2019) (Tempola and Assagaf, 2019) yaitu pada potensi perikanan yang ada di Indonesia dengan algoritma clustering yang digunakan k-means clustering. (Toz, Yücedağ and Erdoğan, 2018) menerapkan *fuzzy image clustering* namun validasi clustering yang digunakan adalah *davies bouldin index*. Kemudian (Khan *et al.*, 2017) meningkatkan cluster k-means dengan melakukan validasi DBI, dimana penerapan DBI dapat di ketahui kualitas cluster terbaik pada dataset yang diterapkan menggunakan algoritma *clustering*.

2. METODE PENELITIAN

Tahap awal dari penelitian ini adalah melakukan pengumpulan data dengan cara survey langsung ke lokasi penelitian yaitu sekolah-sekolah yang berada di kecamatan ternate utara, ternate tengah dan ternate selatan dengan membagikan kuesioner. Selanjutnya hasil kuesioner dilakukan validasi dan reliabilitas pertanyaan. Selanjutnya dilakukan analisis data survey. Sebagaimana langkah-langkah dari penelitian ini ditunjukkan pada Gambar 1.



Gambar 1. Langkah-langkah penelitian

Penelitian ini kuesionernya bersifat tertutup. Adapun variabel operasional yang dilihat dalam survey seperti ditunjukkan pada Tabel 1 yaitu durasi menggunakan jejaring sosial, durasi menonton *youtube*, durasi bermain game online. Jejaring sosial

apa yang digunakan, konten *youtube* apa yang dinonton, game online apa yang dimainkan.

Tabel 1. Operasional Variable Penelitian

Variabel Penelitian	Ukuran	Skala
Pengguna internet pada Anak SD di kota Ternate	Jejaring sosial Youtube Game online	Kategori Kategori kategori

Pada tahap berikutnya adalah melakukan analisis data dari hasil survey pada anak-anak SD di Kota Ternate Provinsi Maluku Utara. Setelah data di analisis maka dilanjutkan dengan melakukan perancangan sistem dengan menerapkan algoritma k-means clustering dengan validasi clustering menggunakan *davies bouldin index*. Dataset yang akan di implementasikan ke sistem adalah berupa rata-rata pengguna aktif dari setiap sekolah pada atribut jejaring sosial, *youtube* dan *game* online. Selanjutnya dilakukan tahapan implementasi sistem dan pengujian sistem sebagai tahapan akhir agar dapat diketahui cluster pengguna internet di sekolah-SD di kota Ternate. Adapun gambaran dari alur penelitian seperti ditunjukkan pada Gambar 1.

2.1. Populasi dan Sampel

Populasi pada penelitian adalah siswa SD di kota Ternate yang berjumlah 16698 (Kemendikbud, 2019) tersebar di 3 kecamatan yaitu Ternate Tengah, Utara dan Selatan. Sedangkan sampelnya adalah Siswa SD kelas 5 di kota Ternate yang berada di Kecamatan Ternate Tengah, Ternate Utara dan Ternate Selatan. Sedangkan untuk sampel dibagi dua bagian yaitu sampel siswa dan sampel sekolah. Dalam penelitian sampel di ambil secara acak atau dikenal dengan teknik *random sampling*.

2.2. K-Means Clustering

Algoritma K-means adalah salah satu bagian dari algoritma clustering, yang mana tergolong pada *unsupervised learning* atau yang sering disebut dengan pembelajaran tanpa guru. *Clustering* berfungsi sebagai pengelompokan obyek kedalam beberapa kelompok berdasarkan kemiripan antar obyek, dimana dalam satu cluster harus berisi obyek yang saling mirip. Clustering ini tanpa menggunakan data pelatihan sebagaimana pada klasifikasi (Budi Santosa dan Ardian Umam, 2018).

Dasar algoritma *K-means* adalah sebagai berikut (Harrington, 2012)

1. Inisialisasi jumlah cluster k
2. Menentukan centroid awal secara random
3. Hitung jarak setiap data ke masing-masing centroid dengan menggunakan konsep kedekatan. Konsep kedekatan ada berbagai macam cara. Namun pada penelitian

menggunakan konsep *euclidean distance*. Sebagaimana ditunjukkan pada Persamaan 1 berikut :

$$d(x_i, y_i) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (1)$$

4. Mengelompokkan setiap data berdasarkan jarak terdekat antara data dengan centroidnya.
5. Tentukan posisi *centroid* baru (k) kembali ke langkah 3 jika posisi *centroid* baru dengan *centroid* lama tidak sama

2.3. Validasi clustering

Cluster yang telah terbentuk perlu untuk dilakukan validasi, tujuannya agar dapat diketahui data tersebut layak dibagi kedalam berapa cluster. Untuk validasi cluster ada beberapa cara salah satunya dengan *davies bouldin index*. Validasi clustering digunakan untuk memutuskan cluster terbaik pada dataset yang diujikan (Viloria and Pineda Lezama, 2019), (Ünlü and Xanthopoulos, 2019) juga memanfaatkan validasi *davies bouldin index* untuk memprediksi jumlah cluster pada sebuah dataset. Metrik Davies-bouldin index (DBI) diperkenalkan oleh David L. Davies dan Donald W. Bouldin (1979) yang tujuannya untuk melakukan evaluasi cluster. Validitas internal yang dilakukan adalah seberapa baik clustering sudah dilakukan dengan menghitung kuantitas dan kriteria atau fitur turunan dari dataset. Adapun persamaan yang digunakan adalah seperti pada Persamaan 2 berikut.

$$DBI = \frac{1}{K} \sum_{i=1}^K \max_{i \neq j} (R_{i,j}) \quad (2)$$

K merupakan jumlah cluster yang digunakan. Sedangkan $R_{i,j}$ adalah rasio perbandingan antara cluster ke- i dan cluster ke- j . Nilainya didapatkan dari komponen kohesi dan separasi. Cluster yang baik adalah yang mempunyai kohesi sekecil mungkin dan separasi sebesar mungkin. $R_{i,j}$ di formulasikan pada Persamaan 3 berikut.

$$R_{i,j} = \frac{SSW_i + SSW_j}{SSB_{i,j}} \quad (3)$$

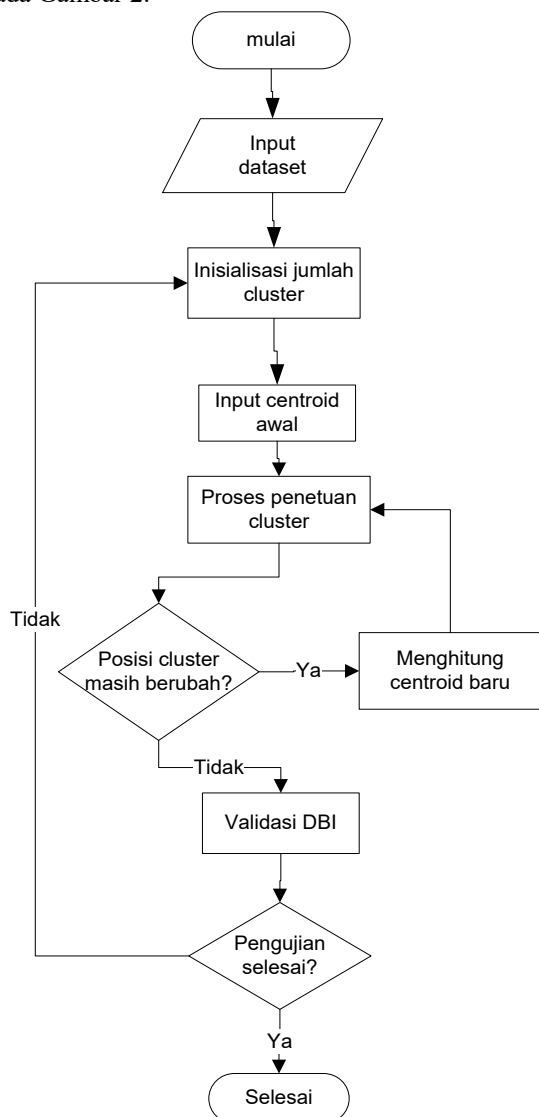
SSB (sum of square between cluster) adalah metrik untuk separasi antara dua cluster, misalnya cluster i dan j , dengan mengukur jarak antara centroid c_i dan c_j seperti pada persamaan 4 berikut.

$$SSB_{i,j} = d(c_i, c_j) \quad (4)$$

Untuk SSW (Sum of Square within cluster) adalah sebagai metrik kohesi dalam suatu cluster ke- i yang mana disajikan pada Persamaan 5 (Davies dan Bouldin, 1975).

$$SSW_i = \frac{1}{m_i} \sum_{j=1}^{m_i} d(x_j, c_i) \quad (5)$$

m_i adalah jumlah data didalam cluster ke- i , sedangkan c_i adalah centroid cluster ke- i . Sedangkan nilai $d()$ dalam persamaan 5 bisa dengan menggunakan *euclidean distance* atau seperti pada Persamaan 1. Adapun langkah-langkah dari penerapan k-means clustering validasi DBI pada penggunaan internet di Kota Ternate ditunjukkan pada Gambar 2.



Gambar 2. Langkah-langkah k-means clustering dan validasi DBI pada penggunaan internet

3. HASIL DAN PEMBAHASAN

Sampel yang dipakai didalam penelitian ini berjumlah 933 siswa. Sampel ini digunakan untuk menganalisis variabel kuesioner. Sedangkan sampel sekolah berjumlah 32 sekolah yang digunakan untuk proses penerapan algoritma clustering k-means dan validasi clustering *davies bouldin index*. jumlah sampel siswa setiap sekolah di tunjukkan pada Tabel 2.

Sebelum diterapkan algoritma clustering k-means dan validasi *davies bouldin index*, maka tahap awal adalah menganalisis hasil survey penggunaan internet pada anak SD di kota Ternate

dari kuesioner yang telah diperoleh kemudian dilakukan proses validitas dan reliabilitas dari pertanyaan.

Tabel 2. Hasil pengumpulan sampel setiap sekolah

No.	Nama Sekolah	Jumlah sampel
1	SD 1	42
2	SD 10	35
3	SD 11	32
4	SD 12	30
5	SD 13	31
6	SD 14	27
7	SD 15	30
8	SD 16	26
9	SD 17	25
10	SD 18	22
11	SD 19	24
12	SD 2	40
13	SD 3	30
14	SD 4	32
15	SD 5	39
16	SD 6	37
17	SD 7	40
18	SD 8	35
19	SD 9	28
20	SD Pertiwi 1	32
21	SD 43	20
22	SD 40	22
23	SD 24	30
24	SD 23	33
25	SD 38	25
26	SD SOA	25
27	SD 42	26
28	SD Muhammadiyah	13
29	SD 32	29
30	SD 28	24
31	SD Islamiyah	30
32	SD 29	22
	Jumlah	933

Hasil Validasi

Berdasarkan hasil validasi yang dilakukan seluruh pertanyaan dinyatakan valid atau mempunyai korelasi hal ini karena r hitung lebih besar dari r tabel sebagaimana hasil uji menggunakan SPSS ditunjukkan pada Gambar 3. Yaitu seluruh nilai total pada setiap pertanyaan lebih besar dari r tabel, dengan nilai df 933.

Correlations				
		x1	x2	x3
x1	Pearson Correlation	1	,503**	,480**
	Sig. (2-tailed)		,000	,000
	N	933	933	933
x2	Pearson Correlation	,503**	1	,612**
	Sig. (2-tailed)	,000		,000
	N	933	933	933
x3	Pearson Correlation	,480**	,612**	1
	Sig. (2-tailed)	,000	,000	
	N	933	933	933
total	Pearson Correlation	,726**	,843**	,899**
	Sig. (2-tailed)	,000	,000	,000
	N	933	933	933

** . Correlation is significant at the 0.01 level (2-tailed).

Gambar 3. Hasil validasi

Hasil reliabilitas

Berdasarkan hasil reliabilitas yang dilakukan pada 3 pertanyaan maka dapat disimpulkan bahwa seluruh pertanyaan handal atau konsisten hal ini terlihat dari

nilai cronebach's alpha 0,742 yang lebih besar dari best practice 0,6. Sebagaimana ditunjukkan pada Gambar 4.

Reliability

Scale: ALL VARIABLES

Case Processing Summary			
		N	%
Cases	Valid	933	100,0
	Excluded ^a	0	,0
	Total	933	100,0

a. Listwise deletion based on all variables in the procedure.

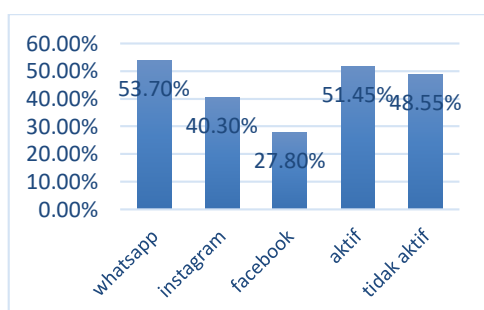
Reliability Statistics

Cronbach's Alpha	N of Items
,742	3

Gambar 4. Hasil uji reliabilitas

Hasil Pengguna Jejaring sosial

Berdasarkan hasil survey pengguna jejaring sosial pada anak-anak SD di kota Ternate dari 933 sampel, didapatkan 53,7 % anak SD di kota Ternate telah memiliki *whatsapp*, 40,3 % memiliki akun *instagram* dan 27,8% memiliki *facebook*. Sebagaimana di tunjukkan pada Gambar 2. Sedangkan yang aktif bersosial media yaitu 51,45 % dan yang tidak aktif bersosial media 48,55 %.



Gambar 2. Graifk Perbandingan Pengguna jejaring sosial

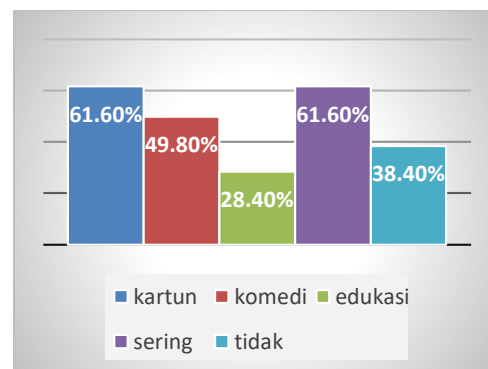
Hasil Pengguna Youtube

Sama halnya pada survey pengguna jejaring sosial, pengguna youtube juga menggunakan sampel yang sama banyak yaitu 933 sampel. Dengan hasil menunjukkan 61,6 % video kartun yang dilihat oleh anak-anak SD di kota ternate, 49,08% melihat konten komedi, sedangkan konten edukasi hanya 28,4 %. Selain itu, ada sekitar 61,6% anak-anak sering membuka youtube, dengan sisanya 38,4% yang jarang membuka youtube. Grafik perbandingan ditunjukkan pada Gambar 3.

Hasil Pencinta Game

Untuk hasil survey pencinta game terdapat 75,34. % anak-anak aktif dalam bermain game namun yang tidak aktif bermain game 24,86%. Sedangkan untuk pada game online sekitar 49,41%

aktif bermain game online sisanya 50,59 tidak aktif bermain game online.



Gambar 3. Grafik perbandingan pengguna youtube

Setelah setiap variabel telah diketahui nilainya masing-masing selanjutnya dilakukan proses clustering dengan k-means dan dilanjutkan dengan validasi *davies bouldin index*. Untuk proses clustering data-data yang diterapkan adalah hasil persentasi pengguna aktif setiap kategori di setiap sekolah sebagaimana di tunjukkan pada Tabel 3 nomor urut sesuai dengan nama sekolah pada Tabel 2.

Tabel 3. Persentasi penggunaan internet setiap sekolah

No.	Jejaring Sosial (%)	Game (%)	Youtube (%)
1	0,67	0,67	0,67
2	0,57	0,71	0,71
3	0,47	0,47	0,47
4	0,53	0,53	0,73
5	0,32	0,32	0,84
6	0,25	0,25	0,71
7	0,67	0,67	0,67
8	0,85	0,85	0,85
9	0,68	0,68	0,68
10	0,23	0,23	0,23
11	0,75	1,00	1,00
12	0,30	0,30	0,30
13	0,47	0,47	0,47
14	0,53	0,53	0,53
15	0,18	0,18	0,18
16	0,27	0,27	0,46
17	0,28	0,28	0,50
18	0,26	0,26	0,46
19	0,27	0,27	0,27
20	0,31	0,31	0,69
21	0,80	0,55	1,00
22	0,59	0,55	0,95
23	0,67	0,47	1,00
24	0,36	0,30	0,97
25	0,96	0,72	1,00
26	0,64	0,68	0,84
27	0,69	0,50	0,92
28	0,77	0,77	0,77
29	0,62	0,90	0,97
30	0,75	0,67	0,17
31	0,80	0,63	0,03
32	0,86	0,59	0,32

Dalam implementasi algoritma clustering k-means dilakukan beberapa tahap pengujian untuk mencari posisi cluster terbaik, untuk dapat diketahui dengan menerapkan *davies bouldin index validation*.

Selanjutnya membandingkan nilai DBI dari setiap pembagian cluster.

Pengujian 2 cluster

Hasil pengujian dengan 2 cluster dengan menerapkan centroid awal secara random didapatkan iterasi sebanyak 3 kali, karena pada iterasi ketiga hasil posisi cluster tidak mengalami perubahan sehingga proses iterasi berhenti. Dari hasil implementasi k-means clustering didapat hasil pembagian cluster 1 ada 16 sekolah begitupun pada cluster 2, sebagaimana ditunjukkan pada Tabel 4. Nilai centroid akhir untuk pembagian 2 cluster yaitu C1 (0,0374; 0,359; 0,456) dan C2 (0,708; 0,678; 0,818).

Tabel 4. Hasil cluster untuk 2 cluster

Cluster 1	Cluster 2
SD 11	SD 1
SD 13	SD 10
SD 14	SD 12
SD 18	SD 15
SD 2	SD 16
SD 3	SD 17
SD 4	SD 19
SD 5	SD 43
SD 6	SD 40
SD 7	SD 24
SD 8	SD 38
SD 9	SD Soa
SD Pertiwi 1	SD 42
SD 23	SD Muhammadiyah
SD 28	SD 32
SD Islamiyah 1	SD 29

Kemudian dilanjutkan dengan melakukan validasi cluster menggunakan davies bouldin index. Untuk mendapatkan nilai DBI maka tahap awal adalah dihitung jarak centroid setiap sekolah menggunakan Persamaan 1. Selanjutnya dihitung nilai *Sum of Square within cluster* (SSW) menggunakan Persamaan 5. Hasil hitung didapat untuk SSW cluster 1 = 0,298. Untuk SSW cluster 2 dihasilkan 0,233. Tahap berikutnya dihitung nilai *sum of square between cluster* (SSB) menggunakan Persamaan 4. dihasilkan nilai SSB di tunjukkan pada Tabel 5. Selanjutnya dicari nilai rasio kedua cluster, hasilnya ditunjukkan pada Tabel 5.

Tabel 5. Hasil nilai SSB dan Rasio

SSB	1	2
1	0	0,586
2	0,586	0
Rasio	1	2
1	0	0,905
2	0,905	0
R Max	0,905	0,905

Berdasarkan pada Tabel 5 nilai *sum of square between cluster* (SSB) dan Rasio pada kedua cluster memiliki nilai sama, hal ini karena jumlah cluster yang terdapat pada cluster 1 dan cluster 2 sama banyak yaitu ada 16 sekolah. Selanjutnya dicari nilai DBI menggunakan Persamaan 2, diperoleh nilai DBI untuk pembagian 2 cluster 0,904.

Pengujian 3 cluster

Proses iterasi pada pembagian 3 cluster terjadi sebanyak 9 kali iterasi, hasil pembagian untuk cluster 1 terdapat 6 sekolah, cluster 2 terdapat 10 sekolah dan cluster 3 terdapat 16 sekolah sebagaimana di tunjukkan pada Tabel 6. Dengan centroid akhir yang dihasilkan adalah C1 (0,647; 0,560; 0,332), C2 (0,260; 0,267; 0,466) dan C3 (0,676; 0,659; 0,858). Dengan nilai DBI yang diperoleh untuk pembagian 3 cluster adalah 0,886.

Tabel 6. Hasil cluster untuk 3 cluster

Cluster 1	Cluster 2	Cluster 3
SD 11	SD 13	SD 1
SD 3	SD 14	SD 10
SD 4	SD 18	SD 12
SD 28	SD 2	SD 15
SD Islamiyah	SD 5	SD 16
SD 29	SD 6	SD 17
	SD 7	SD 19
	SD 8	SD 43
	SD 9	SD 40
	SD Pertiwi 1	SD 24
		SD 23
		SD 38
		SD Soa
		SD 42
		SD
		Muhammadiyah
		SD 32

Pengujian 4 cluster

Hasil pengujian 4 cluster iterasi terjadi sebanyak 8 kali. Dimana pada cluster 1 terdapat 6, sekolah, cluster 2 terdapat 4 sekolah, cluster 3 terdapat 7 sekolah sedangkan cluster 4 terdapat 15 sekolah sebagaimana ditunjukkan pada Tabel 7. Untuk nilai akhir centroid yang dihasilkan untuk C1 (0,647; 0,560; 0,332), C2 (0,245; 0,245; 0,245), C3 (0,283; 0,284; 0,664) dan C4 (0,697; 0,683; 0,851). Sedangkan nilai DBI untuk pengujian 4 cluster yang dihasilkan yaitu 0,773.

Tabel 7. Hasil cluster untuk 4 cluster

Cluster 1	Cluster 2	Cluster 3	Cluster 4
SD 11	SD 18	SD 13	SD 1
SD 3	SD 2	SD 14	SD 10
SD 4	SD 5	SD 6	SD 12
SD 28	SD 9	SD 7	SD 15
SD		SD 8	SD 16
Islamiyah			
SD 29		SD Pertiwi 1	SD 17
		SD 23	SD 19
			SD 43
			SD 40
			SD 24
			SD 38
			SD Soa
			SD 42
			SD
			Muhammadiyah
			SD 32

Pengujian 5 cluster

Hasil pengujian untuk 5 cluster proses iterasi terjadi sebanyak 4 kali iterasi. Dimana pada cluster 1 terdapat 10 sekolah, cluster 2 tidak ada satupun

sekolah yang masuk pada cluster 2, cluster 3 terdapat 3 sekolah, cluster 4 terdapat 4 sekolah dan cluster 5 terdapat 15 sekolah. Dengan nilai centroid akhir yang dihasilkan untuk C1 (0,345; 0,346; 0,612), C2(000; 000; 000), C3 (0,803; 0,630; 0,173), (0,245; 0,245; 0,245), C4 (0,697; 0,683; 0,851). Untuk pengujian 5 cluster proses validasi tidak dilanjutkan hal ini karena pengujian 5 cluster ada cluster yang tidak ada satupun sekolah yang masuk pada cluster tersebut yaitu pada cluster 2. Selain itu, proses pengujian clustering dan validasi DBI berhenti pada pengujian 4 cluster.

Perbandingan Hasil Pengujian Clustering dan validasi DBI

Berdasarkan hasil pengujian penerapan algoritma *clustering* dengan k-means dan validasi *Davies boulding index* (DBI) pada data-data penggunaan internet dikalangan siswa SD di kota Ternate diperoleh nilai DBI untuk pembagian 2 cluster adalah 0,904, pembagian 3 cluster 0,886 dan pembagian 4 cluster 0,773. Sehingga dari ketiga pembagian cluster tersebut nilai Davies bouldin index yang mendekati 0 adalah ketika pengujian 4 cluster yaitu 0,773. Dengan demikian pembagian 4 cluster menjadi posisi cluster terbaik. Selain itu, sekolah-sekolah yang berada pada cluster 1 adalah sekolah dengan persentasi siswanya penggunaan internet pada siswa mereka untuk jejaring sosial adalah pada rentang 47% - 86%, penggunaan persentasi bermain game online 47% - 67% dan penggunaan youtube 3% - 53%. Di cluster 2 sekolah dengan persentasi siswanya menggunakan internet untuk jejaring sosial 18 % - 30%, sama halnya dengan persentasi bermain game online dan menggunakan youtube. Di cluster 3 persentasi menggunakan jejaring sosial 27% - 53%, untuk game online sama juga dengan jejaring sosial dan penggunaan youtube 46% - 97%. Sedangkan cluster 4 persentasi aktif penggunaan jejaring sosial siswa di setiap sekolah adalah 25% - 96%, untuk game online 25% - 100%, sama halnya dengan menggunakan youtube.

4. Kesimpulan Dan Saran

Berdasarkan hasil survey dan penerapan algoritma *clustering k-means* dan validasi *davies bouldin index* pada anak-anak SD di kota Ternate maka diketahui 53,70 % siswa SD menggunakan *whatsapp* dan 51,45 % siswa SD di Kota Ternate aktif menggunakan jejaring sosial. Untuk aktivitas ketika membuka youtube dari 933 siswa SD terdapat 61,60 % menonton film kartun dan 49,80 % menonton komedi. Begitupun dengan yang sering melihat youtube yaitu 61,60%. Sedangkan untuk bermain game online pada siswa SD di Kota Ternate terdapat 49,41 % aktif bermain *game online*. Selain itu, untuk penerapan algoritma clustering k-means dan validasi DBI pada persentasi pengguna aktif untuk tiap kategori jejaring sosial, game online dan youtube di 32 sekolah di kota ternate cluster terbaik

yaitu ketiga pembagian 4 cluster dengan nilai DBI yang diperoleh 0,773.

DAFTAR PUSTAKA

- APRIANTI, W. dan PERMADI, J. (2018) 'K-Means Clustering Untuk Data Kecelakaan Lalu Lintas K-Means', *Jurnal Teknologi Informasi Dan Ilmu Komputer (Jtiik)*, 5(5), Pp. 613-620. Doi: 10.25126/Jtiik2018551113.
- BURHAN, M. S., MAHMUDY, W. F. And DERMAWI, R. (2018) 'Sugeno-Type Fuzzy Inference Optimization With Firefly And K-Means Clustering Algorithms For Rainfall Forecasting', *Journal Of Information Technology And Computer Science*, 3(1), P. 35. Doi: 10.25126/Jitecs.20183134.
- CHANDRA, P. A. (2013) 'Penggunaan Internet Pada Anak-Anak Sekolah Usia 6-12 Tahun Di Surabaya', *Journal Unair*, 1(2). Available At: [Http://Journal.Unair.Ac.Id](http://Journal.Unair.Ac.Id).
- HARRINGTON, P. (2012) *Machine Learning In Action*. 1st Edn. Usa: Manning Publication.
- HUSNI, E. M. dan FATULLOH, A. (2016) 'Kategorisasi Pengguna Internet Di Kalangan Pelajar Sd Dan Smp Menggunakan Metode Twostep Cluster', In *Seminar Nasional Aplikasi Teknologi Informasi*. Yogyakarta, Pp. 6-16. Available At: [Https://Journal.Uii.Ac.Id/Snati/Article/View/6256](https://Journal.Uii.Ac.Id/Snati/Article/View/6256).
- INTERNET SOCIETY.ORG (2017) 'Children And The Internet', *Internet Society.Org*, 26 March. Available At: [Http://www.Internetsociety.Org/Sites/Default/Files/Bpchildrenandtheinternet-](http://www.Internetsociety.Org/Sites/Default/Files/Bpchildrenandtheinternet-).
- INTERNET WORLD STATS.COM (2019) 'Top 20 Countries With The Highest Number Of Internet Users', *Internetworldstats.Com*, 20 March, P. Top 20 Countries With The Highest Number Of Intern. Available At: [Https://www.Internetworldstats.Com/Top20.Htm](https://www.Internetworldstats.Com/Top20.Htm).
- JUMADI D, S, B., Salim Sitompul, O. And Sihombing, P. (2019) 'Enhancement Clustering Evaluation Result Of Davies-Bouldin Index With Determining Initial Centroid Of K-Means Algorithm', In *Journal Of Physics: Conference Series*, P. 012015. Doi: 10.1088/1742-6596/1235/1/012015.
- KEMENDIKBUD (2019) *Data Pokok Pendidikan Dasar Dan Menengah*. Available At: [Https://dapo.Dikdasmen.Kemdikbud.Go.Id/Sp/3/276006](https://dapo.dikdasmen.kemdikbud.go.id/Sp/3/276006).
- KHAN, Z. Et Al. (2017) 'An Improved K-Means Clustering Algorithm Based On An Adaptive Initial Parameter Estimation

- Procedure For Image Segmentation', *International Journal Of Innovative Computing, Information And Control*, 13(5), Pp. 1509–1526. Available At: Scienedirect.Com.
- NCES (2018) *Children's Access To And Use Of The Internet*, Nces.Ed.Gov. Available At: https://nces.ed.gov/programs/coe/pdf/coe_cch.pdf.
- SANTOSA, B. Dan UMAM, A. (2018) *Data Mining And Big Data Analytics. Theories And Implementation Using Python And Apache Spark*. 1st Edn. Yogyakarta: Spreader Media Library.
- TEMPOLA, F. dan ASSAGAF, A. F. (2019) 'Clustering Of Potency Of Shrimp In Indonesia With K-Means Algorithm And Validation Of Davies-Bouldin Index', *Atlantis-Press*, 1(Icst), Pp. 730–733. Doi: 10.2991/Icst-18.2018.148.
- TOZ, G., YÜCEDAĞ, I. dan ERDOĞMUŞ, P. (2018) 'A Fuzzy Image Clustering Method Based On An Improved Backtracking Search Optimization Algorithm With An Inertia Weight Parameter', *Journal Of King Saud University - Computer And Information Sciences*, 31(3), Pp. 295–303. Doi: 10.1016/J.Jksuci.2018.02.011.
- ÜNLÜ, R. dan XANTHOPOULOS, P. (2019) 'Estimating The Number Of Clusters In A Dataset Via Consensus Clustering', *Expert Systems With Applications*. Elsevier Ltd, 125, Pp. 33–39. Doi: 10.1016/J.Eswa.2019.01.074.
- VILORIA, A. dan PINEDA LEZAMA, O. B. (2019) 'Improvements For Determining The Number Of Clusters In K-Means For Innovation Databases In Smes', *Procedia Computer Science*. Elsevier B.V., 151, Pp. 1201–1206. Doi: 10.1016/J.Procs.2019.04.172.