

Penerapan Ensemble K-modes Pada Pengelompokan Kelurahan di Kota Gorontalo Berdasarkan Kecanduan Game Online Remaja

Mohamad Alfiransyah Taufik¹, Novianita Achmad², Siti Nurmardia Abdussamad^{3*}, Djihad Wungguli⁴, Nisky Imansyah Yahya⁵

¹ Mahasiswa Statistika, FMIPA, Universitas Negeri Gorontalo, Gorontalo

^{2, 4, 5} Dosen Matematika, FMIPA, Universitas Negeri Gorontalo, Gorontalo

³ Dosen Statistika, FMIPA, Universitas Negeri Gorontalo, Gorontalo

*sitinurmardia@ung.ac.id

Abstract

The swift advancement of digital technology has resulted in a heightened frequency of online game usage among teenagers, raising concerns about potential addiction.. This study aims to cluster urban villages in Gorontalo City based on the characteristics of online game addiction in adolescents, to support the formulation of more effective preventive policies. The method used is ensemble clustering with K-modes algorithm approach, which is effective for mixed numeric and categorical data. Data were obtained through a survey of adolescents aged 10-24 years in all urban villages, including indicators of lack of attention from close people, self-control, lack of activities, stress or depression, social environment, parenting, length of time playing online games, frequency of playing online games and many favorite online games. The clustering results obtained 3 optimum clusters, where cluster 1 consists of 7 neighborhoods, cluster 2 consists of 17 neighborhoods and cluster 3 consists of 26 neighborhoods. Cluster 1 is a group of neighborhoods with a low risk and addiction level, cluster 2 with a moderate tendency, and cluster 3 with a high tendency. The results of this study are expected to make a significant contribution in mapping areas based on the level of risk and characteristics of online game addiction, which can then be used as a basis for formulating strategic policies and preparing prevention programs tailored to the level of vulnerability of each urban village cluster in Gorontalo City.

Keywords: K-modes; K-medoid; Clustering; Multivariate Analysis; Game Online

Abstrak

Perkembangan teknologi digital yang pesat telah mendorong peningkatan intensitas penggunaan game online di kalangan remaja, yang berpotensi menimbulkan kecanduan. Penelitian ini bertujuan untuk mengelompokkan kelurahan kelurahan di Kota Gorontalo berdasarkan karakteristik kecanduan game online pada remaja, guna mendukung perumusan kebijakan preventif yang lebih efektif. Metode yang digunakan adalah ensemble clustering dengan pendekatan algoritma K-Modes, yang efektif untuk data campuran numerik dan kategorik. Data diperoleh melalui survei terhadap remaja usia 10–24 tahun di seluruh kelurahan, mencakup indikator kurang perhatian dari orang-orang terdekat, kontrol diri, kurang kegiatan, stres atau depresi, lingkungan sosial, pola asuh, lama waktu bermain game online, frekuensi bermain game online dan banyak game online yang digemari. Hasil clustering memperoleh 3 cluster optimum, di mana cluster 1 terdiri dari 7 kelurahan, cluster 2 terdiri dari 17 kelurahan dan cluster 3 terdiri dari 26 kelurahan. Cluster 1 merupakan kelompok kelurahan dengan resiko dan tingkat kecanduan yang cenderung rendah, cluster 2 dengan kecenderungan sedang, dan cluster 3 cenderung tinggi. Hasil penelitian ini diharapkan dapat memberikan kontribusi signifikan dalam pemetaan wilayah berdasarkan tingkat risiko dan karakteristik kecanduan game online, yang selanjutnya dapat dimanfaatkan sebagai landasan dalam

perumusan kebijakan strategis serta penyusunan program pencegahan yang disesuaikan dengan tingkat kerentanan masing-masing cluster kelurahan di Kota Gorontalo.

Kata Kunci: K-modes; K-medoid; Clustering; Analisis Multivariat; Game Online

1. PENDAHULUAN

Perkembangan teknologi informasi telah membawa berbagai kemudahan, salah satunya dalam hal komunikasi dan hiburan. Internet memungkinkan individu untuk berinteraksi tanpa harus bertatap muka serta menyediakan berbagai fitur hiburan, termasuk *game online* (Mais dkk., 2020). Saat ini, *game online* menjadi fenomena global yang digemari oleh berbagai kalangan, mulai dari anak-anak hingga orang dewasa. Tidak hanya sebagai sarana hiburan, *game online* juga dimanfaatkan sebagai media untuk menjalin relasi sosial melalui interaksi daring. Namun, intensitas penggunaan yang tinggi dapat menimbulkan dampak negatif berupa ketergantungan atau kecanduan bermain *game* (Irawan & Siska W, 2021). Menyikapi hal tersebut, sejak tahun 2018, World Health Organization (WHO) secara resmi mengklasifikasikan kecanduan *game online* sebagai gangguan mental dalam International Classification of Diseases (ICD-11). Hal ini menunjukkan bahwa adiksi terhadap *game online* adalah salah satu tantangan kesehatan yang muncul seiring pesatnya perkembangan teknologi informasi.

Pesatnya perkembangan internet mendorong meningkatnya akses terhadap *game online*, yang turut memicu naiknya kasus kecanduan, khususnya di kalangan remaja. Di Indonesia, sebanyak 54,1% remaja usia 15–18 tahun terindikasi kecanduan *game online*, dengan prevalensi lebih tinggi pada laki-laki (32%) dibanding perempuan (22,1%). Pada usia 12–22 tahun, angka kecanduan meningkat menjadi 64,45% pada laki-laki dan 47,85% pada perempuan (Matur dkk., 2021). Di Provinsi Gorontalo, akses internet sangat tinggi, yakni 93,01% di wilayah perkotaan dan 83,60% di pedesaan (BPS, 2023). Selain itu, berdasarkan peta cakupan jaringan Provinsi Gorontalo yang disediakan Nperf menunjukkan bahwa wilayah dengan cakupan jaringan yang paling padat dan kuat terpusat di Kota Gorontalo., menunjukkan tingginya potensi kecanduan *game online* di wilayah tersebut.

Penelitian sebelumnya di Kota Gorontalo oleh Pautina dkk.(2023) menunjukkan bahwa faktor utama penyebab kecanduan *game online* adalah kurangnya perhatian dari orang-orang terdekat (64%), diikuti oleh stres, kurang kegiatan, lingkungan, dan pola asuh. Sementara itu, Fajar dkk.(2024) menemukan bahwa kecanduan *game online* tidak hanya dipengaruhi oleh faktor eksternal seperti lingkungan dan kurangnya perhatian dari orang tua, tetapi juga oleh faktor internal, termasuk stabilitas psikologi, karakter kepribadian, serta kecerdasan emosional.

Berdasarkan penelitian-penelitian sebelumnya, terlihat adanya permasalahan kecanduan *game online* di kalangan remaja. Penelitian ini menggunakan metode

Ensemble *Clustering K-modes* untuk mengelompokkan Kelurahan berdasarkan karakteristik kecanduan *game online* pada remaja. Metode ini memungkinkan identifikasi pola dan faktor-faktor yang memengaruhi kecanduan di tiap kelompok Kelurahan. Dengan memodelkan karakteristik tersebut, diharapkan penelitian ini dapat memberikan rekomendasi yang spesifik bagi setiap kluster Kelurahan dalam upaya pencegahan dan penanganan kecanduan *game online* pada remaja di Kota Gorontalo.

Clustering adalah teknik yang digunakan untuk mengklasifikasikan objek ke dalam kelompok yang relatif homogen, objek-objek dalam setiap *cluster* cenderung mirip satu sama lain dan sangat berbeda dengan objek-objek dari *cluster* lain (Abdussamad, Siti Nurwardia Astutik & Effendi, 2020). Ensemble *clustering* merupakan teknik yang menggabungkan hasil dari beberapa algoritma kluster untuk menghasilkan output akhir yang lebih akurat (Jannah dkk., 2023). Dalam metode ini, data dipisah menjadi numerik dan kategorik; data numerik dianalisis dengan algoritma kluster numerik *k-medoids*, sedangkan data kategorik diproses menggunakan algoritma kluster kategorik *k-modes*. *K-medoids* atau dikenal sebagai Partitioning Around Medoids (PAM), adalah algoritma kluster yang memilih objek data aktual sebagai pusat kluster (medoid). Metode ini termasuk dalam classical partitioning method yang membagi data ke dalam sejumlah kluster berdasarkan kesamaan karakteristik. Jumlah kluster (k) harus ditentukan di awal, dan setiap objek hanya dapat masuk ke satu kluster (Aulanda dkk., 2021). Hasil dari kedua kluster ini kemudian digabung menggunakan algoritma kategorik, seperti *K-modes*, untuk menghasilkan kluster akhir dari data campuran (Shofari dkk., 2024). *K-modes* merupakan pengembangan dari algoritma K-Means yang dirancang khusus untuk mengelompokkan data kategorik. Berbeda dengan K-Means yang menggunakan jarak Euclidean dan nilai rata-rata, *K-modes* menggunakan ukuran jarak berdasarkan kesamaan nilai atribut dan menentukan pusat kluster berdasarkan modus (Badruttamam dkk., 2020). Algoritma ini memperbaiki pusat kluster menggunakan pendekatan berbasis frekuensi untuk meminimalkan jarak total antara data dan pusat klusternya (Yulianton dkk., 2021).

2. METODE PENELITIAN

Data dalam penelitian ini bertipe numerik dan kategorik, sehingga metode analisis yang digunakan adalah algoritma ensemble *clustering* yang memungkinkan untuk menggabungkan hasil *cluster* data numerik dan kategorik. Konsep dari metode ini yaitu menggunakan dua algoritma *clustering* yang berbeda untuk mendapatkan hasil *cluster* data numerik dan kategorik. Dalam penelitian ini digunakan metode ensemble *clustering K-modes*, kemudian untuk memperoleh hasil pengelompokkan data numerik menggunakan metode *cluster K-medoids* dan pengelompokkan data kategorik menggunakan metode analisis *cluster K-modes*.

Data utama dalam penelitian ini merupakan data primer yang dikumpulkan melalui pembagian kuesioner kepada remaja yang menjadi sampel di setiap Kelurahan yang ada

di Kota Gorontalo. Variabel penelitian ini terdiri dari 9 variabel independen yang terbagi menjadi 6 variabel kategorik yaitu Kurang perhatian dari orang-orang terdekat ($X1$), Kontrol diri ($X2$), Kurang Kegiatan ($X3$), Stres atau depresi ($X4$), Lingkungan ($X5$) dan Pola Asuh ($X6$) serta 3 variabel numerik yaitu Lama waktu bermain *game online* ($X7$), Frekuensi bermain *game online* ($X8$) dan Banyak *game online* yang digemari ($X9$).

Populasi dalam penelitian ini adalah 54.637 remaja dengan hasil perhitungan banyak sampel menggunakan slovin yaitu 398 remaja. Teknik penarikan sampel yang digunakan yaitu *purposive sampling* dengan kriteria: (1) Berdomisili tetap di Kota Gorontalo dan (2) Bermain *game online*.

Langkah awal untuk melakukan analisis *cluster* dengan ensemble *k-modes* yaitu dengan memisah data berdasarkan variabel numerik dan kategorik. Kemudian, variabel numerik dianalisis menggunakan algoritma *clustering* numerik yaitu *k-medoids* dan variabel kategorik dianalisis menggunakan algoritma *clustering* kategorik yaitu *k-modes*. Analisis *cluster* data numerik menggunakan *k-medoids* dengan langkah-langkah sebagai berikut (Dinata dkk., 2021):

1. Tentukan k (jumlah *cluster*), dalam penelitian ini menggunakan perhitungan *silhouette coefficient* dengan rumus sebagai berikut (Faujia, Rosi Anisya Setianingsih, Eni Sawitri Pratiwi, 2022):

$$s(i) = \frac{(b_i - a_i)}{\max(b_i, a_i)}$$

dengan a_i adalah rata-rata nilai objek i dengan semua objek dalam *cluster* sendiri dan b_i adalah rata-rata nilai objek i dengan semua objek dalam *cluster* lain.

2. Tentukan titik awal pusat *cluster* (*medoid*) sebanyak k .
3. Hitung jarak *Euclidean* $d(x_i, \mu_j)$ dari setiap objek x_i ke setiap *medoid* μ_j menggunakan rumus berikut:

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

dengan μ_j adalah median (pusat) *cluster* ke- j dan x_i adalah objek data ke- i .

4. Menetapkan objek ke *cluster medoid* terdekat.
5. Ulangi langkah-langkah di atas, dan hitung total deviasi total (S):

$$S = b - a$$

dengan nilai a merupakan total jarak terdekat antara objek dan *medoid* awal, sedangkan b adalah total jarak terdekat antara objek dan *medoid* yang baru.

Apabila nilai S kurang dari nol, maka objek tersebut akan ditukar dengan data lain untuk membentuk k medoid baru.

6. Ulangi langkah ketiga sampai kelima, dan hentikan algoritma jika medoid tidak berubah lagi.

Kemudian untuk analisis data kategorik menggunakan algoritma k -modes dengan langkah-langkah sebagai berikut (Dwiyanti dkk., 2022):

1. Tentukan banyak *cluster* (k) yang akan dibentuk, dalam penelitian ini menggunakan *elbow method* dengan perhitungan *sum of dissimilarity*, rumus sebagai berikut (Pelekis dkk., 2023):

$$SD = \sum_{k=1}^k \sum_{x_i \in C_k} D(x_i, m_k)$$

keterangan:

SD : Nilai *sum of dissimilarity*

k : Jumlah *cluster*

C_k : *Cluster* ke- k

$x_i \in C_k$: Anggota *cluster* ke- i dalam *cluster* ke- k

m_k : Pusat dari *cluster* k

$D(x_i, m_k)$: *Dissimilarity* (ketidaksamaan) antara data x_i dan pusat *cluster* m_k .

2. Tentukan modus awal pusat *cluster* (centroid) sebanyak k .
3. Menghitung jarak setiap objek ke masing-masing centroid. Jarak ditentukan melalui perhitungan simple matching *dissimilarity* atau ukuran ketidakmiripan sederhana dengan rumus sebagai berikut:

$$d(A, B) = \sum_{j=1}^n \delta(x_{1j}, x_{2j})$$

dengan

$$\delta(x_{1j}, x_{2j}) = \begin{cases} 0, & \text{jika } x_{1j} = x_{2j} \\ 1, & \text{jika } x_{1j} \neq x_{2j} \end{cases}$$

dengan $X1_j$ merupakan atribut dari variabel ke- j pada objek A1 dan $X2_j$ adalah atribut dari variabel ke- j pada objek A2.

4. Menempatkan objek sesuai dengan jarak terdekat ke centroid. Ukuran ketidakmiripan sederhana dapat digunakan untuk menunjukkan *cluster* yang paling dekat dengan data tersebut.

5. Setelah semua data ditempatkan ke *cluster* terdekat, langkah berikutnya adalah memperbarui centroid setiap *cluster* dengan menggunakan modus dari setiap variabel anggota *cluster* yang ada.
6. Menghitung kembali jarak setiap data ke seluruh centroid baru dengan menggunakan ukuran ketidakmiripan sederhana seperti pada langkah ketiga.
7. Melakukan pengulangan pada langkah kelima dan keenam sampai tidak ada data yang berpindah ke *cluster* lain.

Selanjutnya dilakukan pengelompokkan dari gabungan hasil *cluster* numerik dan kategorik menggunakan algoritma ensemble *k-modes* untuk memperoleh hasil *cluster* akhir, langkah-langkah yang digunakan sama dengan tahapan algoritma *k-modes* untuk data kategorik diatas. Data yang digunakan merupakan penggabungan dari hasil analisis *cluster* numerik dan kategorik yang nantinya menjadi variabel baru.

3. HASIL DAN PEMBAHASAN

Karakteristik perilaku bermain *game online* di Kota Gorontalo digambarkan melalui analisis statistik deskriptif pada masing-masing variabel numerik dan kategorik.

Tabel 1. Statistik deskriptif

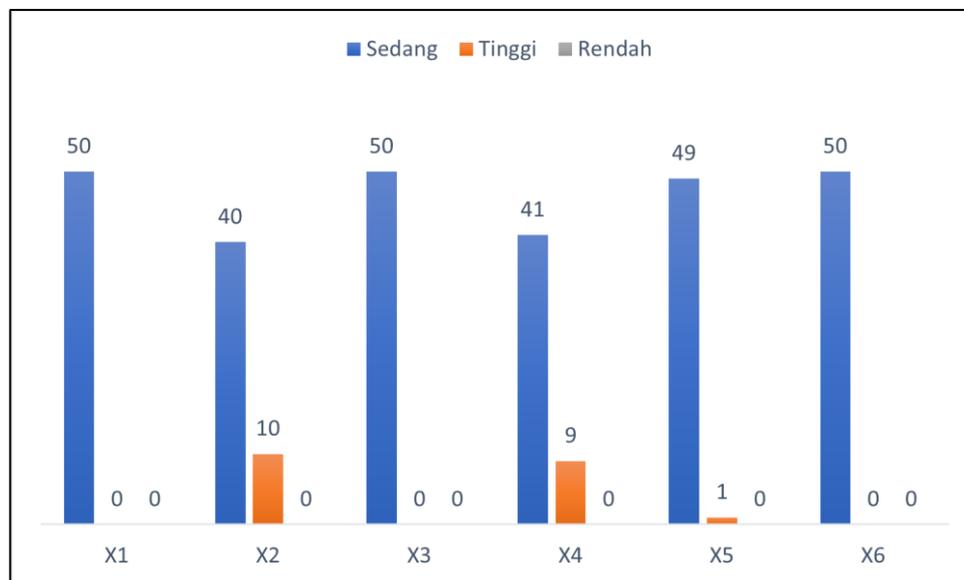
Variabel	Mean	Min	Max
Lama waktu bermain <i>game online</i> (X7)	3,24	1,91	5,20
Frekuensi bermain <i>game online</i> (X8)	2,32	1,14	3,56
Banyak <i>game online</i> yang digemari (X8)	2,26	1,33	3,75

Berdasarkan tabel 2 terlihat bahwa durasi atau rata-rata lama waktu bermain *game online* (X7) remaja di setiap kelurahan di kota Gorontalo adalah 3,24 jam atau 3 jam 14 menit. Pada frekuensi bermain *game online* (X8), rata-rata tiap kelurahan memiliki remaja dengan frekuensi bermain lebih dari 2 kali setiap harinya, yaitu sebesar 2,32 kali bermain. Rata-rata banyak *game online* yang digemari (X9) oleh remaja di setiap kelurahan adalah lebih dari 2 *game* dengan nilai 2,26.

Lama waktu bermain *game online* (X7) terendah terdapat pada kelurahan Pulubala yaitu 1,91 jam atau selama 1 jam 54 menit, sedangkan yang tertinggi terdapat pada kelurahan Botu yaitu 5,20 jam atau 5 jam 12 menit. Frekuensi bermain *game online* (X8) terendah terdapat pada kelurahan Dulalowo yaitu 1,14 kali atau hampir hanya satu kali per hari, sedangkan yang tertinggi terdapat pada kelurahan Padebuolo yaitu 3,56 kali atau hampir 4 kali per hari. Banyak *game online* yang digemari (X9) oleh remaja paling sedikit terdapat pada kelurahan Tomulabutao yaitu sebanyak 1,33 *game* atau hampir hanya 1 *game* saja yang digemari, sedangkan yang terbanyak terdapat pada kelurahan Pilolodaa yaitu sebanyak 3,75 *game* atau hampir 4 *game* yang digemari.

Variabel bertipe kategorik yang digunakan pada penelitian ini yaitu Kurang perhatian dari orang-orang terdekat (X1), Kontrol diri (X2), Kurang kegiatan (X3), Stres atau

Depresi ($X4$), Lingkungan sosial ($X5$) dan Pola asuh ($X6$). Karakteristik tingkat kecanduan *game online* di Kota Gorontalo berdasarkan variabel kategorik ditampilkan pada diagram batang pada gambar berikut.



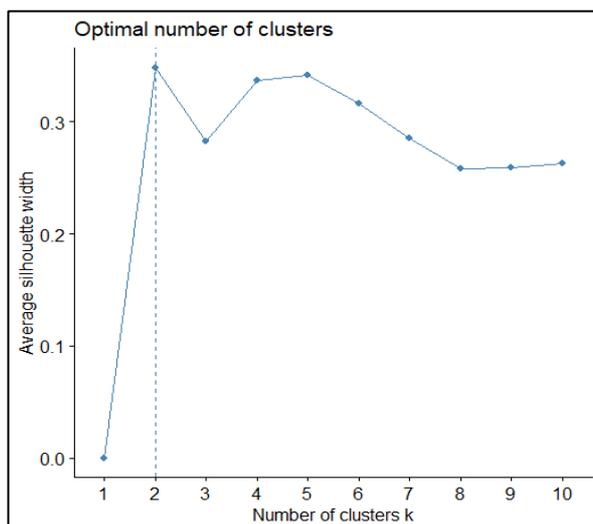
Gambar 1. Statistik deskriptif variabel kategorik

Diagram batang pada Gambar 1 menunjukkan distribusi kategori Sedang dan Tinggi untuk variabel kategorik $X1$ hingga $X6$, dapat disimpulkan bahwa secara umum mayoritas data berada pada kategori Sedang. Variabel $X1$, $X3$, dan $X6$ untuk seluruh kelurahan di Kota Gorontalo sepenuhnya berada dalam kategori Sedang. Ini menunjukkan bahwa seluruh kelurahan pada ketiga variabel tersebut belum mencapai kategori Tinggi. Variabel $X5$ juga didominasi oleh kategori Sedang sebanyak 49 kelurahan, dengan hanya satu kelurahan yang termasuk dalam kategori Tinggi yaitu kelurahan Tamalate. Hal ini menunjukkan bahwa variabel ini hampir seragam pada tingkat sedang. Untuk variabel $X2$ dan $X4$, terdapat sedikit variasi, masing-masing memiliki 10 dan 9 kelurahan yang sudah termasuk dalam kategori Tinggi, sedangkan sisanya masuk pada kategori Sedang. Meskipun demikian, kategori Sedang tetap mendominasi pada kedua variabel ini. Secara keseluruhan, kategori rendah tidak muncul pada semua variabel dan tidak terdapat variabel yang menunjukkan dominasi kategori Tinggi, yang mengindikasikan bahwa sebagian besar data masih berada dalam kondisi menengah. Secara keseluruhan, data kategorik ini menunjukkan bahwa resiko dan tingkat kecanduan *game online* pada setiap kelurahan di Kota Gorontalo cenderung masih berada pada tingkat sedang, dan belum banyak yang menunjukkan gejala kecanduan berat.

3.1 Pengelompokkan data numerik dengan algoritma *K-medoids*

Pengelompokkan data numerik pada penelitian ini menggunakan algoritma *clustering K-medoids* untuk mengelompokkan kelurahan di Kota Gorontalo berdasarkan variabel

Lama waktu bermain *game online* ($X7$), Frekuensi bermain *game online* ($X8$) dan Banyak *game online* yang digemari ($X9$). *Silhouette coefficient* (SC) digunakan untuk menentukan jumlah *cluster* optimal.



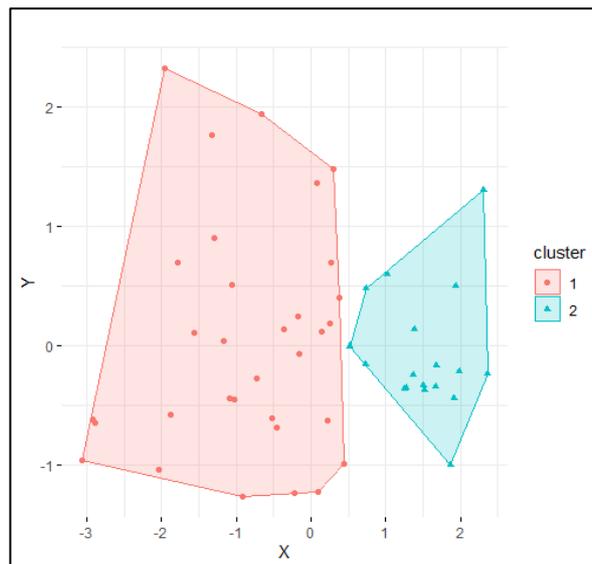
Gambar 2. Plot SC

Berdasarkan gambar 2 terlihat bahwa jumlah *cluster* optimal adalah 2, karena memberikan rata-rata *silhouette* tertinggi yang menandakan bahwa dua *cluster* sudah cukup memisahkan data dengan baik secara internal (kompak) dan eksternal (terpisah dari *cluster* lain). Sehingga dalam pengelompokan data numerik menggunakan *K-medoids* akan menetapkan jumlah *cluster* sebanyak 2 kelompok. Hasil pengelompokan ditampilkan pada tabel berikut.

Tabel 2. Anggota kelompok hasil *cluster K-medoids*

Cluster	Kelurahan
1	Biawao, Biawu, Botu, Buladu, Buliide, Bulotadaa Timur, Dembe I, Donggala, Dulalowo Timur, Dulomo Selatan, Dulomo Utara, Heledulaa Selatan, Heledulaa Utara, Huangobotu, Ipilo, Leato Selatan, Libuo, Liluwo, Limba B, Limba U I, Molosipat W, Padebuolo, Pilolodaa, Siendeng, Talumolo, Tamalate, Tenda, Tomulabutao Selatan, Tuladenggi, Wongkaditi Barat, Wongkaditi Timur dan Wumialo.
2	Bugis, Bulotadaa Barat, Dembe II, Dembe Jaya, Dulalowo, Leato Utara, Lekobalo, Limba U II, Molosipat U, Moodu, Paguyaman, Pohe, Pulubala, Tanggikiki, Tanjung Kramat, Tapa, Tenilo dan Tomulabutao.

Berdasarkan tabel 2 terlihat bahwa pengelompokan tersebut menghasilkan sebanyak 32 kelurahan dikelompokkan pada *cluster* 1 dan sebanyak 18 kelurahan dikelompokkan pada *cluster* 2. Hasil tersebut divisualisasikan pada bidang dua dimensi untuk melihat bagaimana data terbagi ke dalam masing-masing *cluster* berdasarkan kemiripan karakteristik.

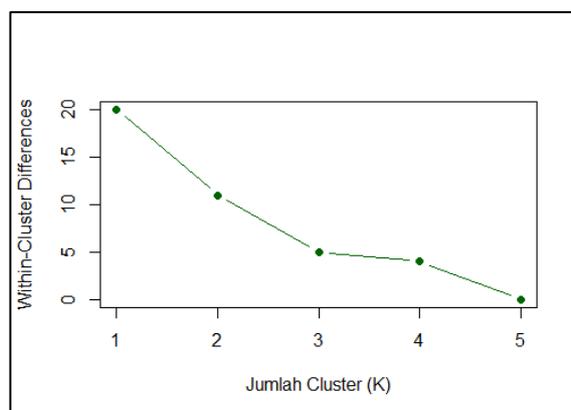


Gambar 3. Plot klasterisasi

Berdasarkan hasil klasterisasi yang divisualisasikan dalam gambar 3, data berhasil dikelompokkan ke dalam dua *cluster* yang terpisah dengan jelas. *cluster* pertama terdiri dari data yang tersebar lebih luas yang berarti bahwa *cluster* ini mencakup kelompok dengan karakteristik yang lebih bervariasi. Sebaliknya, *cluster* kedua memiliki persebaran data yang lebih padat dan terkonsentrasi, *cluster* ini menggambarkan kelompok data yang relatif lebih homogen. Secara keseluruhan, hasil klasterisasi ini menunjukkan bahwa metode pengelompokan yang digunakan mampu memisahkan data secara efektif ke dalam dua kelompok yang berbeda, baik dari segi posisi maupun pola persebarannya.

3.2 Pengelompokan data kategorik dengan algoritma *K-modes*

Pengelompokan data kategorik menggunakan algoritma *K-modes* untuk mengelompokkan Kelurahan di Kota Gorontalo berdasarkan variabel Kurang perhatian dari orang-orang terdekat ($X1$), Kontrol diri ($X2$), Kurang Kegiatan ($X3$), Stres atau depresi ($X4$), Lingkungan ($X5$) dan Pola Asuh ($X6$). Metode *elbow* dengan menghitung *sum of dissimilarity* digunakan untuk menentukan jumlah *cluster* optimal.

Gambar 4. Plot *elbow* untuk *K-modes*

Berdasarkan gambar 4 terlihat bahwa titik tekuk terlihat berada pada $k = 3$ yang menjadi indikator kuat bahwa jumlah *cluster* ideal untuk klasterisasi data kategorik ini adalah tiga. Sehingga dalam pengelompokkan data kategorik menggunakan *K-modes* akan menetapkan jumlah *cluster* sebanyak 3 kelompok. Hasil pengelompokkan ditampilkan dalam tabel berikut.

Tabel 3. Anggota kelompok hasil *cluster K-modes*

<i>Cluster</i>	Kelurahan
1	Dulalowo Timur, Huangobotu, Lekobalo, Libuo, Limba B, Limba U I, Pulubala, Tanjung Kramat dan Wongkaditi Timur.
2	Buliide, Bulotadaa Timur, Dembe I, Dembe II, Dembe Jaya, Donggala, Dulalowo, Dulomo Selatan, Dulomo Utara, Heledulaa Selatan, Leato Selatan, Leato Utara, Liluwo, Molosipat U, Padebuolo, Pilolodaa, Siendeng, Talumolo, Tamalate dan Tenda
3	Biawao, Biawu, Botu, Bugis, Buladu, Bulotadaa Barat, Heledulaa Utara, Ipilo, Limba U II, Molosipat W, Moodu, Paguyaman, Pohe, Tanggikiki, Tapa, Tenilo, Tomulabutao, Tomulabutao Selatan, Tuladenggi, Wongkaditi Barat dan Wumialo

Berdasarkan tabel 3 terlihat bahwa pengelompokkan tersebut menghasilkan sebanyak 9 kelurahan dikelompokkan pada *cluster* 1, sebanyak 20 kelurahan dikelompokkan pada *cluster* 2 dan sebanyak 21 kelurahan dikelompokkan pada *cluster* 3.

3.3 Pengelompokkan data campuran dengan algoritma Ensemble *K-modes*

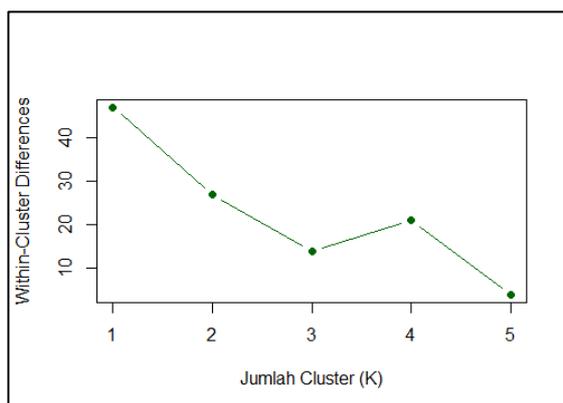
Hasil pengelompokkan variabel numerik dan kategorik memperoleh hasil *cluster* optimal masing-masing 2 dan 3 *cluster*. Selanjutnya dilakukan pengelompokkan dari gabungan hasil *cluster* numerik dan kategorik menggunakan algoritma ensemble *k-modes* untuk memperoleh hasil *cluster* akhir. Data yang digunakan merupakan penggabungan dari hasil analisis *cluster* numerik dan kategorik yang nantinya menjadi variabel baru. Hasil *cluster* numerik dituliskan sebagai variabel baru $X1$ dan hasil *cluster* kategorik dituliskan sebagai variabel baru $X2$ seperti dalam tabel berikut.

Tabel 4. Data hasil *cluster* numerik ($X1$) dan kategorik ($X2$)

Kelurahan	$X1$	$X2$	Kelurahan	$X1$	$X2$
Biawao	1	3	Limba U I	1	3
Biawu	1	3	Limba U II	2	1
Botu	1	3	Liluwo	1	2
Bugis	2	3	Molosipat U	2	2
Buladu	1	3	Molosipat W	1	3
Buliide	1	2	Moodu	2	3
Bulotadaa Barat	2	3	Padebuolo	1	2
Bulotadaa Timur	1	2	Paguyaman	2	3
Dembe I	1	2	Pilolodaa	1	2

Kelurahan	X1	X2	Kelurahan	X1	X2
Dembe II	2	2	Pohe	2	3
Dembe Jaya	2	2	Pulubala	2	1
Donggala	1	2	Siendeng	1	2
Dulalowo	2	2	Talumolo	1	2
Dulalowo Timur	1	1	Tamalate	1	2
Dulalowo Selatan	1	2	Tanggikiki	2	3
Dulalowo Utara	1	2	Tanjung Kramat	2	1
Heledulaa Selatan	1	2	Tapa	2	3
Heledulaa Utara	1	3	Tenda	1	2
Huangobotu	1	1	Tenilo	2	3
Ipilo	1	3	Tomulabutao	2	3
Leato Selatan	1	2	Tomulabutao Selatan	1	3
Leato Utara	2	2	Tuladenggi	1	3
Lekobalo	2	1	Wongkaditi Barat	1	3
Libuo	1	1	Wongkaditi Timur	1	1
Limba B	1	1	Wumialo	1	3

Jumlah *cluster* optimal untuk ensemble *k-modes* pada penelitian ini ditentukan melalui metode *elbow* dengan perhitungan *sum of dissimilarity*.



Gambar 5. Plot *elbow* untuk ensemble *K-modes*

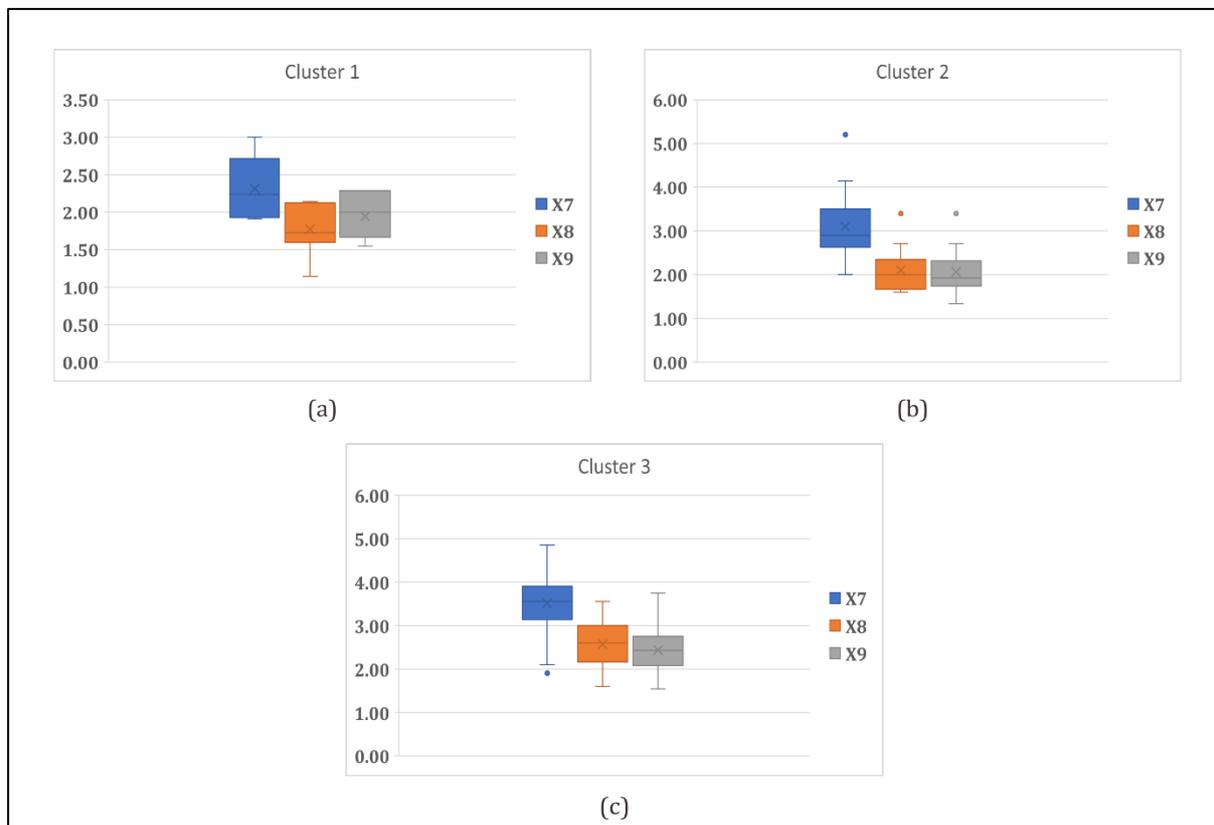
Berdasarkan gambar 5 terlihat bahwa titik tekuk terlihat berada pada $k = 3$ yang menjadi indikator bahwa jumlah *cluster* optimal untuk data kategorik ini adalah tiga. Sehingga dalam pengelompokan data campuran menggunakan ensemble *k-modes* akan menetapkan jumlah *cluster* sebanyak 3 kelompok. Hasil pengelompokan ditampilkan dalam tabel berikut.

Tabel 5. Anggota kelompok hasil *cluster* Ensemble *K-modes*

Cluster	Kelurahan
1	Dembe II, Dembe Jaya, Dulalowo, Leato Utara, Lekobalo, Molosipat U dan Pulubala.
2	Biawu, Botu, Bugis, Buladu, Bulotadaa Barat, Limba U II, Moodu, Paguyaman, Pohe, Tanggikiki, Tanjung Kramat, Tapa, Tenilo, Tomulabutao,

Cluster	Kelurahan
	Tomulabutao Selatan, Tuladenggi, WumialoLeato Utara, Liluwo, Molosipat U, Padebuolo, Pilolodaa, Siendeng, Talumolo, Tamalate dan Tenda.
3	Biawao, Buliide, Bulotadaa Timur, Dembe I, Donggala, Dulalowo Timur, Dulomo Selatan, Dulomo Utara, Heledulaa Selatan, Heledulaa Utara, Huangobotu, Ipilo, Leato Selatan, Libuo, Liluwo, Limba B, Limba U I, Molosipat W, Padebuolo, Pilolodaa, Siendeng, Talumolo, Tamalate, Tenda, Wongkaditi Barat dan Wongkaditi Timur.

Berdasarkan tabel 5 terlihat bahwa pengelompokkan tersebut menghasilkan sebanyak 7 kelurahan dikelompokkan pada *cluster* 1, sebanyak 17 kelurahan dikelompokkan pada *cluster* 2 dan sebanyak 26 kelurahan dikelompokkan pada *cluster* 3.

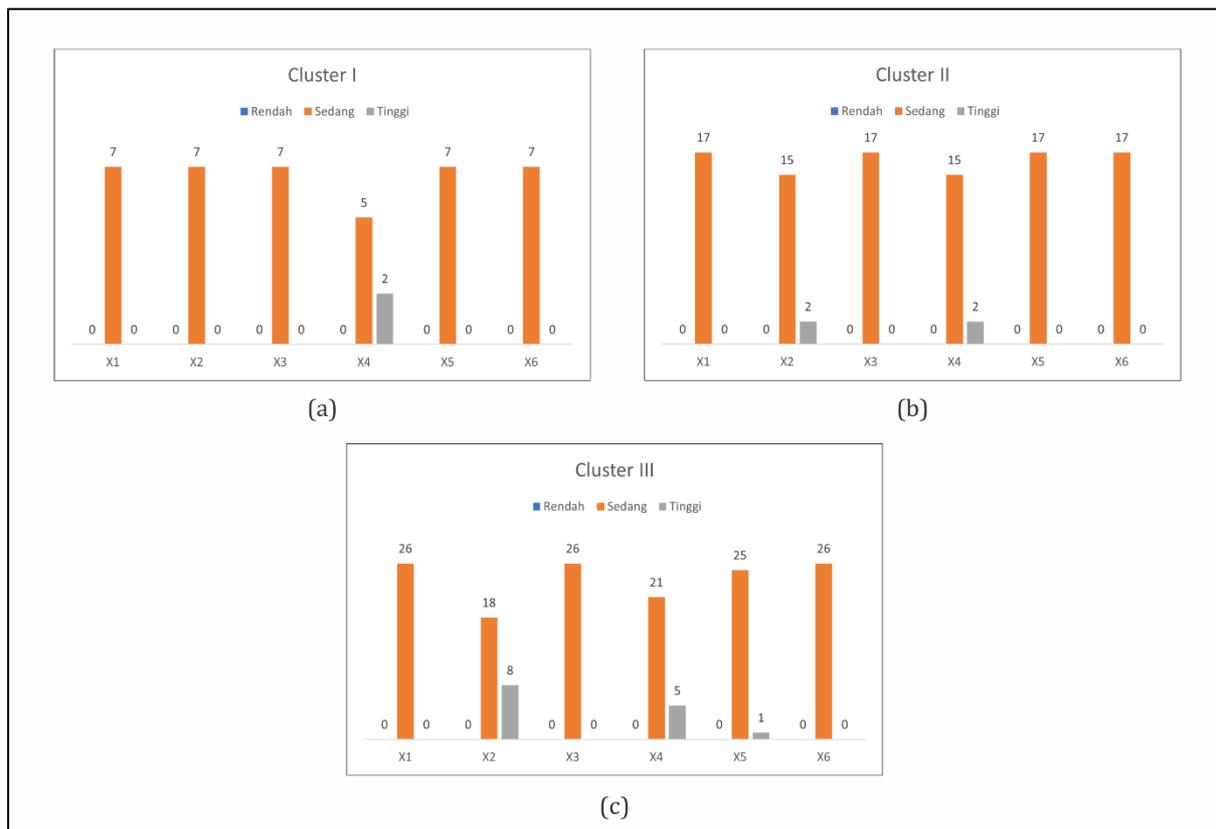


Gambar 6. Boxplot variabel numerik tiap *cluster*

Berdasarkan Gambar 6 terlihat bahwa *cluster* 1 menunjukkan tingkat keterlibatan paling rendah dalam aktivitas bermain *game online*. Rata-rata lama waktu bermain *game online* (*X7*) hanya 2,32 jam per hari dengan nilai terendah adalah 1,91 jam dan tertinggi adalah 3 jam, frekuensi bermain (*X8*) mendekati 2 kali sehari (1,77) dengan distribusi yang lebih lebar diatas median menunjukkan bahwa ada sebagian kecil individu yang bermain *game* dengan frekuensi tinggi, dan banyak *game* yang digemari (*X9*) rata-rata 1,95 (sekitar 2 *game*).

Cluster 2 memiliki tingkat keterlibatan yang sedang. Rata-rata lama waktu bermain *game online* (*X7*) adalah 3,1 jam per hari dengan nilai terendah adalah 2 jam dan tertinggi adalah 5,2 jam yang terindikasi sebagai outlier, hal ini menandakan bahwa terdapat kelurahan pada *cluster* ini dengan durasi bermain yang ekstrim. Frekuensi bermain (*X8*) rata-rata adalah 2 kali sehari (2,10) dengan frekuensi tertinggi yaitu 3,4 (hampir 4 kali sehari) yang juga menjadi nilai ekstrim, dan rata-rata banyak *game* yang digemari (*X9*) sekitar 2 *game* (2,06) dengan nilai tertinggi yaitu 3,4 atau hampir 4 *game online* yang digemari terindikasi sebagai nilai ekstrim.

Cluster 3 merupakan kelompok dengan keterlibatan tertinggi dalam bermain *game online*. Rata-rata lama waktu bermain *game online* (*X7*) mencapai 3,56 jam per hari dengan durasi terendah adalah 1,91 jam yang terindikasi sebagai outlier, hal ini menandakan bahwa ada kelurahan yang memiliki durasi bermain sangat rendah pada *cluster* ini dengan durasi tertinggi adalah 4,86 jam per hari. Frekuensi bermain (*X8*) rata-rata 2–3 kali sehari (2,58) dengan frekuensi tertinggi adalah 3,56 atau hampir 4 kali sehari. Rata-rata banyak *game* yang digemari (*X9*) lebih banyak yaitu sekitar 2–3 *game* (2,44) dengan nilai tertinggi yaitu 3,75 atau jika dibukatkan yaitu sebanyak 4 *game* yang digemari. Dengan demikian, dapat ditentukan bahwa *cluster 1* merupakan kelompok dengan karakteristik kecanduan *game online* yang cenderung rendah, *cluster 2* sedang dan *cluster 3* merupakan kelompok dengan kecenderungan tinggi.



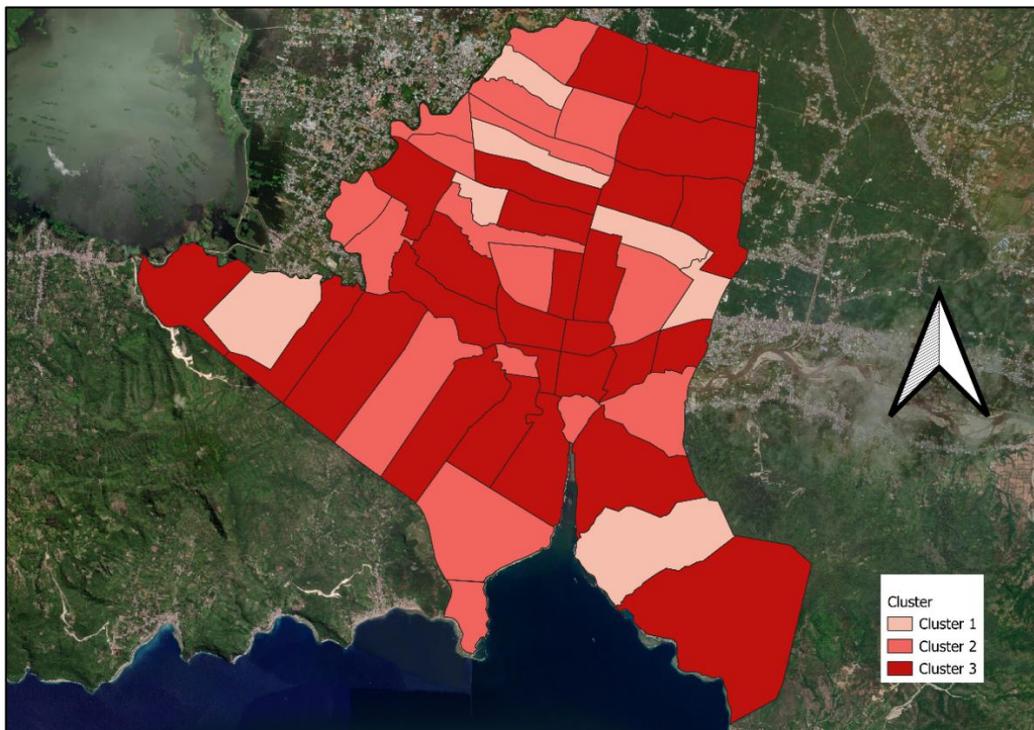
Gambar 7. Plot variabel kategorik tiap *cluster*

Berdasarkan gambar 7 terlihat bahwa *cluster* 1 menunjukkan karakteristik yang cenderung homogen. Seluruh variabel dalam *cluster* ini didominasi oleh kategori sedang, sedangkan kategori tinggi hanya muncul pada variabel stres (X_4) sebanyak 2 kelurahan. Pola ini mencerminkan bahwa kelurahan-kelurahan yang tergolong dalam *cluster* 1 secara umum mengalami resiko dan tingkat kecanduan *game online* yang sedang, tanpa adanya kecenderungan kuat ke arah kategori rendah ataupun tinggi, meskipun ada sedikit kecenderungan stres yang lebih tinggi pada 2 kelurahan.

Cluster 2 juga didominasi oleh kategori sedang, namun mulai muncul variasi. Terdapat kategori tinggi pada faktor kontrol diri (X_2) dan stres (X_4) yaitu sebanyak dua kelurahan. Variasi tersebut hanya sebagian kecil, yang mengindikasikan bahwa wilayah dalam *cluster* 2 memiliki resiko dan tingkat kecanduan *game online* yang sedang.

Cluster 3 memperlihatkan distribusi yang lebih beragam. Meski kategori sedang masih dominan, terdapat kategori tinggi yang cukup banyak pada faktor kontrol diri (X_2), dengan sedikit kategori tinggi pada faktor stres (X_4) serta lingkungan (X_5). Ini mengindikasikan bahwa beberapa kelurahan dalam *cluster* 3 memiliki remaja dengan kontrol diri yang baik, akan tetapi secara keseluruhan masih memiliki resiko dan tingkat kecanduan *game online* yang sedang.

Berdasarkan hasil visualisasi distribusi kategori, dapat dilihat bahwa ketiga *cluster* cenderung memiliki karakteristik yang serupa, yaitu didominasi oleh kategori sedang pada keenam variabel mulai dari variabel kurang perhatian orang-orang terdekat, kontrol diri, kurang kegiatan, stres, lingkungan, dan pola asuh. Ini menunjukkan bahwa variabel-variabel tersebut tidak memiliki perbedaan yang signifikan di ketiga *cluster*.



Gambar 7. Pemetaan hasil *clustering*

Secara visual hasil dari pengelompokan data campuran ditampilkan di atas peta wilayah Kota Gorontalo pada gambar 8. Pemetaan hasil *clustering* terhadap tingkat kecanduan *game online* menunjukkan adanya tiga kategori wilayah yang masing-masing tergolong ke dalam *cluster* 1 (berwarna merah pudar), *cluster* 2 (berwarna merah jambu), dan *cluster* 3 (berwarna merah pekat). Wilayah-wilayah yang termasuk dalam *cluster* 1 tampak tersebar secara terbatas pada peta. Sebaliknya, *cluster* 2 mencakup area yang cukup luas. Sementara itu, *cluster* 3 yang ditandai dengan warna merah pekat terlihat paling mendominasi secara visual di peta. Hasil *clustering* dalam penelitian ini sejalan dengan hasil studi oleh Pautina dkk. (2023) yang menekankan pentingnya faktor-faktor sosial seperti perhatian keluarga dan kondisi psikologis dalam memengaruhi tingkat kecanduan game online. Meskipun pendekatan yang digunakan berbeda, kesamaan fokus lokasi dan objek kajian menguatkan bahwa kecanduan game online merupakan isu yang kompleks dan dipengaruhi oleh berbagai faktor sosial serta lingkungan, yang penting untuk dipetakan secara spasial guna mendukung intervensi yang lebih tepat sasaran. Oleh karena itu, hasil pemetaan wilayah risiko ini dapat menjadi dasar bagi penyusunan program intervensi berbasis wilayah yang tidak hanya mempertimbangkan tingkat kecanduan, tetapi juga kondisi sosial dan emosional masyarakat di tiap kelurahan.

4. SIMPULAN

Hasil analisis dan pembahasan terhadap karakteristik kecanduan *game online* remaja pada 50 kelurahan di Kota Gorontalo menghasilkan kesimpulan sebagai berikut:

1. Berdasarkan implementasi ensemble *clustering k-modes* dapat diketahui bahwa terdapat 3 *cluster* optimum yang terbentuk dengan masing-masing *cluster* terdiri dari 7 kelurahan pada *cluster* 1, 17 kelurahan pada *cluster* 2 dan 26 kelurahan pada *cluster* 3.
2. Berdasarkan hasil *cluster* yang terbentuk, pertama melalui statistik deskriptif data numerik dapat diketahui bahwa *cluster* 1 merupakan kelompok dengan resiko dan kecanduan *game online* yang cenderung rendah, *cluster* 2 sedang dan *cluster* 3 merupakan kelompok dengan kecenderungan tinggi. Kemudian, melalui statistik deskriptif data kategorik dapat diketahui bahwa *cluster* 1 secara umum mengalami resiko dan tingkat kecanduan *game online* yang sedang, tanpa adanya kecenderungan kuat ke arah kategori rendah ataupun tinggi, meskipun ada sedikit kecenderungan stres yang lebih tinggi pada 2 kelurahan. *Cluster* 2 juga didominasi oleh kategori sedang, namun mulai muncul variasi tetapi hanya sebagian kecil, yang secara keseluruhan mengindikasikan bahwa wilayah dalam *cluster* 2 memiliki resiko dan tingkat kecanduan *game online* yang sedang. *Cluster* 3 memiliki remaja dengan kontrol diri yang baik, akan tetapi secara keseluruhan masih memiliki resiko dan tingkat kecanduan *game online* yang sedang.
3. Berdasarkan pemetaan hasil *clustering* pada peta wilayah Kota Gorontalo, dapat diketahui bahwa wilayah-wilayah yang termasuk dalam *cluster* 1 tampak tersebar secara terbatas pada peta. Sebaliknya, *cluster* 2 mencakup area yang cukup luas. Sementara itu, *cluster* 3 terlihat paling mendominasi secara visual di peta.

6. REKOMENDASI

Dari hasil analisis serta kesimpulan penelitian, rekomendasi saran yang diajukan untuk penelitian selanjutnya yaitu pada penerapan algoritma ensemble *clustering k-modes* agar menggunakan data kategorik yang lebih seimbang dalam distribusi kategori tingkat kecanduan, baik dari aspek jumlah responden maupun representasi dari tiap tingkat (rendah, sedang, tinggi). Hal ini penting untuk menghasilkan *cluster* yang lebih variatif dan mendeskripsikan perbedaan karakteristik secara lebih signifikan.

7. REFERENSI

- Abdussamad, Siti Nurmardia Astutik, S., & Effendi, A. (2020). Evaluation of Implementation Context Based Clustering in Fuzzy Geographically Weighted Clustering-Particle Swarm Optimization Algorithm. *Jurnal EECCIS*, 14.
- Aulanda, L., Windarto, A. P., & Okprana, H. (2021). Pengelompokan Pembiayaan Nasabah Klaim Asuransi Pengguna Kendaraan Bermotor dengan Metode K-Medoids. *TIN: Terapan Informatika Nusantara*, 2.
- Badruttamam, A., Sudarno, & Maruddani, D. A. I. (2020). Penerapan Analisis Kluster K-Modes Dengan Validasi Davies Bouldin Index dalam Menentukan Karakteristik Kanal Youtube Di Indonesia (Studi Kasus: 250 Kanal YouTube Indonesia Teratas Menurut Socialblade). *Jurnal Gaussian*, 9.
- Dinata, R. K., Retno, S., & Hasdyna, N. (2021). Minimization of the Number of Iterations in K-Medoids Clustering with Purity Algorithm. *International Information & Engineering Technology Association*, 35.
- Dwiyamti, S. N., Nisa, K., Sutrisno, A., & Herawati, N. (2022). Analisis Kluster untuk Data Kategorik Menggunakan Metode K-Modes (Studi Kasus: Data Pasien COVID-19 di RSUD Dr. H. Abdul Moeloek Provinsi Lampung). *Jurnal Siger Matematika*, 03.
- Fajar, M., Masyhuri, M., & Muda, Y. (2024). Kecanduan Game Online pada Remaja. *Journal of Education Research*, 5.
- Faujia, Rosi Anisya Setianingsih, Eni Sawitri Pratiwi, H. (2022). Analisis Kluster K-Means Dan Agglomerative Nesting Pada Indikator Stunting Balita Di Indonesia. *Seminar Nasional Official Statistics*.
- Irawan, S., & Siska W, D. (2021). Faktor-Faktor Yang Mempengaruhi Kecanduan Game Online Peserta Didik. *Jurnal Konseling Gusjigang*, 07.
- Jannah, B., Utami, I. T., & Hakim, A. R. (2023). Metode ensemble Robust Clustering Using Links(Rock) Untuk Pengelompokan Perguruan Tinggi Swasta (PTS) Di Kota Semarang. *Jurnal Gaussian*, 12.
- Mais, F. R., Rompas, S. S. J., & Gannika, L. (2020). Kecanduan Game Online Dengan Insomnia Pada Remaja. *Jurnal Keperawatan*, 8(2), 18. <https://doi.org/10.35790/jkp.v8i2.32318>
- Matur, Y. P., Simon, M., & Ndorang, T. (2021). Hubungan Kecanduan Game Online Dengan Kualitas Tidur Pada Remaja SMA Negeri Di Kota Ruteng. *Wawasan Kesehatan*, 6.
- Pautina, M. R., Tuasikal, J. M. S., & Siregar, I. K. (2023). Deskripsi Faktor-Faktor Yang Mempengaruhi Siswa Kecanduan Game Online Di SMP Negeri 1 Kota Gorontalo. *Superior Education Journal*, 1.
- Pelekis, S., Pipergias, A., Karakolis, E., Mouzakitis, S., Santori, F., Ghoreishi, M., & Askounis, D.

(2023). Targeted demand response for flexible energy communities using clustering techniques. *Sustainable Energy, Grids and Networks*, 36.

Shofari, M. R., Soesanto, O., & Kartini, D. (2024). Quick Robust Clustering Using Links (Qrock) Untuk Pengelompokan Desa Kabupaten Banjar. *RAGAM: Journal of Statistics and Its Application*, 3.

Yulianton, H., Sutanto, F. A., & Mulyani, S. (2021). Pengelompokan Mahasiswa Berbasis Categorical Variables Menggunakan Metode K-Modes Clustering. *Proceeding SENDIU*.