

PENGELOMPOKAN GRAPHICS PROCESSING UNIT MENGGUNAKAN K-MEANS CLUSTERING

¹Qoharudin Al-GhonyuRohimullah

²Agung Nilogiri

Jurusan Teknik Informatika, Fakultas Teknik, Universitas Muhammadiyah Jember

¹Email : pqohar@gmail.com

ABSTRAK

Pada saat ini teknologi berkembang dengan cepat dan banyak teknologi yang mulai bermunculan termasuk dalam dunia komputer, salah satu komponen yang terpenting dalam sebuah komputer adalah graphics processing unit (GPU) yang digunakan dalam mengolah gambar. Fungsi gpu yang lain adalah untuk bermain sebuah game, pada tahun 2015 dirt rally adalah sebuah game yang banyak di mainkan karena memiliki grafik yang baik dan dapat menggunakan resolusi tinggi yaitu 3840 X 2160 pixel (4K). Dengan banyaknya pilihan gpu saat ini akan menyulitkan setiap orang untuk membeli gpu apa yang sesuai dengan pengguna tersebut, karena setiap seri gpu memiliki performa yang berbeda. Dengan adanya metode k-means clustering diharapkan dapat membantu pengguna untuk mengelompokkan seri gpu berdasarkan rata-rata fps dan minimum fps pada game dirt rally. Algoritma davies bouldin-index merupakan algoritma yang efisien dalam menentukan validasi pada suatu metode pengelompokan, yang nantinya akan menghasilkan kelompok mana yang terbaik dari semua pengelompokan yang dilakukan. Hasil pengelompokan dari 101 data gpu pada tahun 2015 sampai 2017 yang dikelompokkan dari 2 kelompok sampai 10 kelompok didapatkan pengelompokan yang memiliki nilai DBI yang terbaik adalah pengelompokan dengan 2 kelompok yang memiliki nilai DBI 0,28429, dengan data dalam cluster 1 sebanyak 23 data dan data dalam cluster 2 sebanyak 78 data.

Kata kunci : *Graphics processing unit (GPU), K-Means , Davies Bouldin-index*

ABSTRACT

At this time technology is developing rapidly and many technologies are beginning to emerge, including in the computer world, one of the most important components in a computer is the graphics processing unit (GPU) used in processing images. Another function of the gpu is to play a game, in 2015 dirt rally is a game that is widely played because it has good graphics and can use high resolution of 3840 X 2160 pixels (4K). With so many choices the current GPU will make it difficult for everyone to buy a GPU that suits the user, because each GPU series has a different performance. The existence of the k-means clustering method is expected to help users to group the GPU series based on the average fps and minimum fps in the dirt rally game. Bouldin-index davies algorithm is an efficient algorithm in determining the validation of a grouping method, which in turn will produce which group is the best of all groupings performed. The results of grouping from 101 GPU data in 2015 to 2017 grouped from 2 groups to 10 groups obtained the grouping that has the best DBI value is grouping with 2 groups that have a DBI value of 0.28429, with data in cluster 1 totaling 23 data and data in cluster 2 there are 78 data.

Keywords: *Graphics processing unit (GPU), K-Means, Davies Bouldin-index*

1. Pendahuluan

1.1 Latar Belakang

Pada saat ini teknologi berkembang dengan sangat cepat dan juga semakin banyak teknologi canggih yang mulai bermunculan. Perkembangan teknologi ini pastinya juga berkaitan dengan perkembangan teknologi komputer yang semakin lama semakin banyak teknologi di bidang komputer yang di ciptakan,

karena komputer diciptakan untuk membantu mempermudah pekerjaan manusia dan juga sebagai sarana hiburan. Komputer yang kita gunakan didalamnya memiliki beberapa komponen agar komputer kita bisa digunakan untuk mempermudah pekerjaan kita. Salah satu komponen terpenting dalam pengolahan grafis yang ada di komputer kita adalah *Graphics Processing unit (GPU)*. GPU adalah sebuah

hardware komputer yang berfungsi untuk menampilkan hasil dari pemrosesan yang dilakukan oleh CPU. Pada saat ini GPU lebih banyak digunakan dalam bermain sebuah *game* dan digunakan pada pengolahan grafis, GPU saat ini sangat banyak kita temukan di pasaran mulai dari *custom card* dari beberapa *vendor* ataupun GPU *founders edition*.

Salah satu *game* yang memiliki *graphical gameplay* yang spektakuler sehingga membutuhkan GPU atau kartu grafis yang memiliki performa yang baik agar *game* yang dimainkan memiliki *frame rate* yang tinggi adalah DIRT RALLY. DIRT RALLY adalah sebuah yang ber-*genre racing* atau balap, dan dapat digunakan sampai dengan resolusi 3840 X 2160 (4K).

Pada saat ini terdapat banyak pilihan seri dari sebuah GPU, banyaknya seri tersebut akan menyulitkan pengguna untuk memilih GPU yang ingin dibeli sesuai dengan kebutuhan, setiap seri pada sebuah GPU memiliki performa yang berbeda. Pada penelitian ini GPU dikelompokkan berdasarkan performa dalam FPS dan dilakukan uji performa seri GPU dengan menggunakan *game* DIRT RALLY, *game* DIRT RALLY adalah satu-satunya *game* di pengujian *website hardwareluxx* yang tetap digunakan dari tahun 2015 – 2018 dengan pengujian sistem yang sama mulai dari *processor* (CPU), *random access memory* (RAM), penyimpanan, *Motherboard* dan *setting game* yang sama akan tetapi yang berbeda adalah seri GPU tersebut agar beberapa seri dari GPU tersebut bisa dibandingkan dari segi performanya. Hasil dari pengujian tersebut menghasilkan rata – rata FPS dan *minimum FPS* yang didapatkan dari setiap seri GPU, rata – rata FPS dan *minimum FPS* tersebut yang biasanya menyebabkan efek *stutter* dimonitor yang kemudian merusak pengalaman bermain *game* [8].

K-Means *clustering* dapat digunakan sebagai alternative metode pengelompokan GPU, sehingga dapat menghasilkan informasi kelompok GPU yang terdapat dalam masing-masing *cluster* yang memiliki kemiripan performanya. Dengan mengetahui kelompok dari GPU yang memiliki kemiripan tersebut

diharapkan pengguna dapat terbantu dalam mengambil keputusan saat memilih dan membeli sebuah GPU yang sesuai dengan performa dalam satu *cluster* GPU tersebut.

1.2 Rumusan Masalah

Rumusan masalah yang dijadikan objek dalam penelitian ini berdasarkan latar belakang yang telah dipaparkan adalah :

- 1) Berapa jumlah *cluster* optimum pada pengelompokan GPU menggunakan metode algoritma *K-Means*?
- 2) Berapa jumlah GPU dalam masing-masing *cluster* yang terpilih?

1.3 Batasan Masalah

Untuk lebih memfokuskan pengerjaan penulisan ini di tetapkan batasan antara lain:

- 1) GPU yang digunakan memiliki seri GPU antara lain: AMD (R9, R9 FURY, R9 FURY X, RX 480, RX 580, RX VEGA 56 dan RX VEGA 64) dan NVIDIA GTX (770, 780, 780 TI, 970, 980, 980 TI, 1060, 1070, 1070 TI, 1080, 1080 TI, TITAN (KEPLER), TITAN X (MAXWELL), TITAN X (PASCAL) dan TITAN V).
- 2) Data didapatkan dari *website hardwareluxx.de* menggunakan pengujian *game* DIRT RALLY dengan resolusi (3840 X 2160) *pixels* atau biasa di sebut dengan resolusi 4K, *setting* yang digunakan di dalam *game* adalah *ultra* dengan 1 X *anti-aliasing* dan 1 X *antisoptic filtering*.
- 3) Data yang digunakan dalam pengujian ini adalah nama GPU, *average FPS* dan *minimum FPS*
- 4) Penentuan *cluster optimum* berdasarkan teknik *Davies-Buildin Index*
- 5) Skema uji *cluster* pada penelitian ini dimulai dari 2 sampai dengan 10 *cluster*.
- 6) *Tools* yang digunakan adalah Rapid Miner.

1.4 Tujuan

Adapun tujuan dari penulisan untuk tugas akhir ini adalah sebagai berikut:

- 1) Mencari jumlah *cluster* optimum pada pengelompokan GPU menggunakan metode *K-Means*

- 2) Mengetahui jumlah GPU dalam masing-masing kelompok yang terbentuk.

1.5 Manfaat

Adapun manfaat dari penulisan tugas akhir tentang GPU ini adalah sebagai berikut:

2. Tinjauan Pustaka

2.1 Graphics Processing Unit (GPU)

GPU adalah singkatan dari *graphics processing unit*, GPU biasanya juga disebut sebagai kartu grafis. Setiap komputer atau laptop akan menggunakan GPU untuk melakukan tugas menampilkan gambar, video, animasi 2D dan animasi 3D. GPU melakukan tugas perhitungan matematika cepat dan membebaskan CPU untuk melakukan hal-hal lain. CPU menggunakan beberapa *core* yang berfokus pada pemrosesan *serial*, sedangkan GPU memiliki ribuan *core* lebih kecil yang digunakan untuk melakukan tugas *multi-tasking*. GPU terdapat 2 jenis yaitu GPU diskrit dan juga *integrated* GPU. GPU diskrit adalah sebuah GPU yang memiliki VRAM sendiri dan juga bentuknya terpisah dari CPU, sedangkan *integrated* GPU adalah GPU yang terletak pada *processor* sehingga membutuhkan RAM yang ada di *processor* sebagai pengganti dari VRAM [5].

Nvidia dan Radeon adalah dua pioneer GPU diskrit di dunia yang memiliki beragam teknologi canggih pada kartu grafisnya. Umumnya sistem komputer yang telah menggunakan kartu grafis discrete atau external akan memiliki performa grafis yang lebih baik dari GPU terintegrasi. Tetapi akibatnya tenaga yang diperlukan semakin besar, serta panas yang dihasilkan juga semakin tinggi, karena itu keberadaan GPU terintegrasi seperti milik Intel masih dibutuhkan pada sistem yang berskala kecil ataupun mobile untuk menghemat tenaga serta mengurangi produksi panas berlebih. Masing-masing vendor memiliki fitur tersendiri yang telah ditanamkan pada kartu grafis mereka yang akan menentukan performanya [9].

2.2 Data Mining

Data *mining* adalah proses menemukan pola yang menarik dan pengetahuan dari data yang berjumlah besar [4]. Data mining adalah aktivitas yang menggambarkan sebuah proses

- 1) Sebagai metode alternatif dalam pengelompokan GPU
- 2) Memberikan pilihan GPU lain yang performanya sama atau mirip, jika GPU yang ingin dibeli tidak terdapat di toko.

analisis yang terjadi secara iteratif pada database yang besar, dengan tujuan mengekstrak informasi dan *knowledge* yang akurat dan berpotensi berguna untuk *knowledge workers* yang berhubungan dengan pengambilan keputusan dan pemecahan masalah [10].

Aktivitas data *mining* dapat dipisahkan menjadi 6, berdasarkan tujuan dari analisis yaitu [6]: Estimasi, Deskripsi, Prediksi, Klasifikasi, *Clustering*, Asosiasi.

2.3 K-Means

K-means merupakan metode klusterisasi yang paling terkenal dan banyak digunakan diberbagai bidang karena sederhana, mudah diimplementasikan, memiliki kemampuan untuk mengkluster data yang besar dan kompleksitas waktunya linear. *K-means* merupakan metode pengklusteran secara *partitioning* yang memisahkan data ke dalam kelompok yang berbeda. Dengan *partitioning* secara iterasi, *K-Means* mampu meminimalkan rata - rata jarak setiap data ke klusternya [11].

Dasar algoritma *K-means* adalah sebagai berikut:

1. Tentukan nilai k sebagai jumlah kluster yang ingin dibentuk.
2. Bangkitkan k *centroid* (titik pusat kluster) awal secara *random/acak*, kemudian untuk menghitung *centroid cluster* k-i berikutnya, digunakan rumus sebagai berikut:

$$v = \sqrt{\frac{\sum_{i=1}^n x_i}{n}} \dots \dots \dots (2.1)$$

Dimana:

- i : 1,2,3,...n
- v : *centroid* pada *cluster*
- x_i : objek ke-i
- n : banyaknya objek/jumlah objek yang menjadi anggota *cluster*

3. Hitung jarak setiap data ke masing-masing *centroid* menggunakan rumus.

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \dots \dots \dots (2.2)$$

Ket :

- d_{ij} : Jarak objek antara objek x dan y
- n : Jumlah Atribut
- x_i : Objek Data
- y_i : Data Cluster

4. Kelompokkan setiap data berdasarkan jarak terdekat antara data dengan *centroid*-nya (C).
5. Kembali ke langkah 3 jika posisi *centroid* baru dengan *centroid* lama tidak sama. [1]

2.4 Davies-Bouldin Index

Davies-Bouldin Index merupakan salah satu metode yang digunakan untuk mengukur validitas cluster pada suatu metode pengelompokan, kohesi didefinisikan sebagai jumlah dari kedekatan data terhadap titik pusat cluster dari cluster yang diikuti, Evaluasi menggunakan Davies Bouldin Index ini memiliki skema evaluasi internal cluster, dimana baik atau tidaknya hasil cluster dilihat dari kuantitas dan kedekatan antar data hasil cluster [2].

Sum of square within cluster (SSW) merupakan persamaan yang digunakan untuk mengetahui matrik kohesi dalam sebuah cluster ke-i yang dirumuskan sebagai berikut:

$$SSW_i = \frac{1}{m_i} \sum_{j=i}^{m_i} d(x_j, c_i) \dots \dots \dots (2.3)$$

Dari persamaan tersebut, m_i merupakan jumlah data dalam cluster ke-i, c_i adalah *centroid cluster* ke-i, dan $d()$ merupakan jarak setiap data ke *centroid* yang dihitung menggunakan jarak *euclidean*.

Sum of square between cluster (SSB) merupakan persamaan yang digunakan untuk mengetahui separasi antar cluster yang dihitung menggunakan persamaan:

$$SSB_{i,j} = d(c_i, c_j) \dots \dots \dots (2.4)$$

Setelah nilai kohesi dan separasi diperoleh, kemudian dilakukan pengukuran rasio (R_{ij}) untuk mengetahui nilai perbandingan antara cluster ke-i dan cluster ke-j. Cluster yang baik adalah cluster yang memiliki nilai kohesi sekecil mungkin dan separasi yang

sebesar mungkin. Nilai rasio dihitung menggunakan persamaan sebagai berikut:

$$R_{i,j} = \frac{SSW_i + SSW_j}{SSB_{i,j}} \dots \dots \dots (2.5)$$

Nilai rasio yang diperoleh tersebut digunakan untuk mencari nilai *davies-bouldin index (DBI)* dari persamaan berikut:

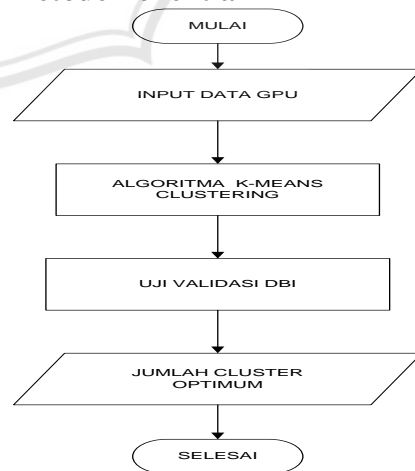
$$DBI = \frac{1}{k} \sum_{i=1}^k \max_{i \neq j} (R_{i,j}) \dots \dots \dots (2.6)$$

Dari persamaan tersebut, k merupakan jumlah cluster yang digunakan. Semakin kecil nilai DBI yang diperoleh (non-negatif ≥ 0), maka semakin baik cluster yang diperoleh dari pengelompokan *K-means* yang digunakan [7]

2.5 Rapid Miner

Menurut Lee dan Juan [3] Rapid Miner (Yale) adalah “perangkat lunak open source untuk knowledge discovery dan data mining.” Rapid Miner memiliki kurang lebih 400 prosedur (operator) data mining, termasuk operator untuk masukan, output, data preprocessing dan visualisasi. Ribuan aplikasi data mining yang telah dikembangkan menggunakan Rapid Miner banyak digunakan di dunia bisnis maupun penelitian. Beberapa fitur dari Rapid Miner, antara lain: Berlisensi gratis (open source), multiplatform karena diprogram dalam bahasa Java dan internal data berbasis XML sehingga memudahkan pertukaran data eksperimen. dibaca oleh Rapid Miner untuk menjalankan analisis secara otomatis.

3. Metode Penelitian



Gambar 3 Flowchart Proses Pengelompokan GPU.

Langkah-langkah dalam menentukan cluster terbaik adalah sebagai berikut :

1. Menginputkan data GPU, data yang digunakan dalam perhitungan adalah GPU name, average fps dan minimum fps.
2. Data GPU tersebut di cluster menggunakan algoritma K-Means sehingga menghasilkan data yang ter-cluster.
3. Data yang sudah ter-cluster kemudian dilakukan pencarian nilai Davies-Bouldin Index, sehingga nilai DBI digunakan untuk menentukan cluster terbaik.
4. Langkah clustering dengan K-Mean dan pencarian nilai DBI dilakukan dengan menggunakan 2 sampai dengan 10 cluster.
5. Nilai DBI yang terkecil dari 2 sampai 10 cluster menunjukkan bahwa pengelompokan data tersebut yang paling optimum.

3.1 Pengelompokan Data

Data yang digunakan dalam penelitian ini adalah data dari beberapa GPU dengan berbagai custom card dari beberapa vendor. Data GPU yang digunakan dalam penelitian ini adalah data yang menggunakan test bench yang memiliki spesifikasi yang sama akan tetapi berbeda GPU saja. Pengumpulan data dalam penelitian ini di dapatkan dari situs resmi di internet yaitu dari hardwarelux dengan situs yang berasal dari Jerman.

3.2 Proses Clustering

Data dalam penelitian ini merupakan data GPU dari NVIDIA dan juga AMD yang berjumlah sebanyak 101 data. Pengelompokan tersebut berdasarkan average fps dan juga minimum fps dari game DIRT RALLY yang menggunakan resolusi (3840 X 2160) atau biasa di sebut dengan resolusi 4K. Pada pengelompokan ini penentuan cluster validasinya menggunakan teknik Davies-Bouldin Index. Proses clustering pada penelitian ini dimulai dari 2 cluster sampai dengan 10 cluster.

3.3 Proses Davies-Bouldin Index

Penelitian ini menentukan cluster optimumnya menggunakan teknik Davies-Bouldin Index. teknik Davies-Bouldin Index digunakan untuk mencari nilai DBI yang terkecil dari masing-masing cluster, dan cluster yang memiliki nilai DBI yang terendah adalah cluster yang paling optimum dari semua cluster yang ada.

3.4 Hasil Clustering dan Perhitungan DBI

Pengujian yang dilakukan dari 101 data GPU dengan menggunakan bantuan aplikasi RapidMiner Studio 9.5.001. Parameter yang digunakan menggunakan clustering dari 2 sampai dengan 10 cluster, didapatkan nilai pada tabel dibawah ini.

No	Jumlah Cluster	Nilai DBI
1	2 cluster	0,284
2	3 cluster	0,34
3	4 cluster	0,415
4	5 cluster	0,348
5	6 cluster	0,443
6	7 cluster	0,426
7	8 cluster	0,369
8	9 cluster	0,413
9	10 cluster	0,414

Hasil dari pengujian yang dilakukan menggunakan 101 data GPU berdasarkan tabel diatas, pengelompokan menggunakan 2 cluster memiliki nilai DBI yang lebih kecil dari 2 sampai dengan 10 cluster yang membuktikan pengelompokan menggunakan 2 cluster memiliki nilai yang lebih optimum dari pada pengelompokan lain.

CENTROID (C)	1	2
Rata – Rata FPS	153,065217	62,6884615
FPS Terendah	4	4
	123	54,8974359

Tabel diatas menunjukkan centroid dari pengelompokan menggunakan 2 cluster dengan menggunakan atribut rata – rata FPS dan juga FPS terendah, rata – rata FPS dan FPS terendah didapatkan dari pengujian yang dilakukan di website hardwarelux. Nilai centroid untuk rata – rata FPS pada pengelompokan 2 cluster di cluster 1 adalah 153,0652174 dan nilai centroid

FPS terendah pada pengelompokan 2 cluster di cluster 1 adalah 123, nilai *centroid* untuk rata – rata FPS pada pengelompokan 2 cluster di cluster 2 adalah 62,68846154 dan nilai *centroid* FPS terendah pada pengelompokan 2 cluster di cluster 2 adalah 54,8974359.

4. Kesimpulan dan Saran

4.1 Kesimpulan

Berdasarkan hasil yang diperoleh, didapat kesimpulan sebagai berikut :

1. *Cluster optimum* yang didapatkan dari penelitian ini adalah pengelompokan menggunakan 2 cluster dengan nilai DBI terkecil yaitu 0,284.
2. Jumlah data pada cluster 1 sebanyak 23 data dan jumlah data pada cluster 2 sebanyak 78 data pada pengelompokan menggunakan 2 cluster.
3. Rentang rata – rata FPS tiap data pada cluster 1 diantara 142,5 FPS sampai dengan 172,3 FPS sedangkan untuk nilai rata – rata FPS tiap data pada cluster 2 diantara 12,1 FPS sampai dengan 93 FPS.
4. Rentang FPS *minimum* pada cluster 1 adalah 119 sampai dengan 133 sedangkan untuk FPS *minimum* pada cluster 2 adalah 9 sampai dengan 83
5. Rentang rata – rata FPS dan nilai *minimum* FPS pada cluster 1 lebih tinggi dari pada cluster 2.

4.2 Saran

1. Penilaian selanjutnya dapat menggunakan pengujian *game* yang lebih baru dan fitur *game* yang lebih baru.
2. Dapat dilanjutkan dengan menggunakan pengujian *benchmark sintetis* yang lebih baru.

5. Daftar Pustaka

- [1] Agusta, Y. 2007. *K-Means* penerapan permasalahan dan metode terkait. Jurnal Sistem dan Informatika, Vol. 3, pp 47-60.
- [2] Bates, A., Kalita, J. 2016. *Counting clusters in twitter posts. Proceedings of the 2nd International Conference on Information*

Technology for Competitive Strategies, pp. 85

- [3] Finn, L., Santana, J. 2010. *Data Mining: Meramalkan Bisnis Perusahaan*. Jakarta: Elex Media Komputindo.
- [4] Han, J., Kamber, M. 2011 *Data Mining Concepts and Techniques*. San Fransisco: Morgan Kaufmann.
- [5] Harding, S. 2019. What is a GPU? A Basic Definition of Graphics Cards. Toms Hardware. <https://www.tomshardware.com/reviews/gpu-graphics-card-definition,5742.html> diakses tanggal 25 juni 2020.
- [6] Larose, D. T. 2005 *Discovering Knowledge In Data*. America: Wiley.
- [7] Prasetyo, E. 2016. *Data Mining Mengolah Data menjadi Informasi dengan Matlab*. Yogyakarta: Andi.
- [8] Ramadhan, R. A. 2015. *Apa itu AMD FRTC (Frame Rate Target Control)*. Jagat Review. <http://www.jagatreview.com/2015/07/apa-itu-amd-frtc-frame-rate-target-control/3/> diakses tanggal 20 maret 2020.
- [9] Santoso, J. D. 2019. *Analisis Password Cracking Menggunakan GPU Process*. Fakultas Ilmu Komputer, Universitas AMIKOM, Vol. 3, No. 1.1, pp 143-150.
- [10] Vercellis, C. 2009. *Business Intelligence: Data Mining and Optimization for Decision Making*. Italy: Wiley.
- [11] Warowaruntu, M, N, V., Amin, M, F. 2018. Penerapan Metode K-Means Pemetaan Calon Penerima Jamkesda. Sekolah Tinggi Manajemen Informatika dan Komputer. Banjarbaru. Vol. 5, no 2, pp 190-200.