

## **PENERAPAN *K-MEDOIDS CLUSTERING* PADA EKSPOR BUAH-BUAHAN MENURUT NEGARA TUJUAN**

**Aris Saputri<sup>1</sup>, Hidayatullah<sup>2\*</sup>, Ari Dermawan<sup>3</sup>**

<sup>1</sup>Sistem Informasi, STMIK Royal Kisaran

<sup>2</sup>Sistem Komputer, STMIK Royal Kisaran

<sup>3</sup>Manajemen Informatika, STMIK Royal Kisaran

*\*email* : dayatscorpio2@gmail.com

**Abstrack:** The aim of the study is to group fruit exports according to the country of destination. The research data used came from the Indonesian Central Statistics Agency with the url <https://www.bps.go.id/> for the category of fruit exports by destination country. The computer science technique used is to utilize K-Medoids clustering data lamination. The results of the study are expected to provide information to the government about the mapping results in the form of clusters of the destination countries for the number of fruit exports. This needs to be done to review the process of fruit exports to destination countries, bearing in mind the results of the export of these fruits have the potential to improve the Indonesian economy.

**Keywords:** Data mining; Klastering; K-Medoids; Fruit Export; Destination Country.

**Abstrak:**Tujuan dari penelitian adalah untuk mengelompok kan ekspor buah-buahan menurut negara tujuan. Data penelitian yang digunakan berasal dari Badan Pusat Statistik Indonesia dengan kategori ekspor buah-buahan menurut negara tujuan. Teknik ilmu komputer yang digunakan adalah dengan memanfaatkan data mining klastering K-Medoids. Hasil penelitian diharapkan dapat memberikan informasi kepada pemerintah tentang hasil pemetaan berupa cluster terhadap Negara tujuan untuk jumlah ekspor buah-buahan. Hal ini perlu dilakukan untuk meninjau ulang proses ekspor buah buah kenegara tujuan mengingat hasil ekspor buah-buah tersebut berpotensi untuk meningkatkan perekonomian Indonesia.

**Kata Kunci:** Data mining; Klastering; K-Medoids; Ekspor Buah; Negara Tujuan.

## **PENDAHULUAN**

Ekspor merupakan mengirim barang dari suatu Negara ke negara lainnya dengan tujuan untuk mendapatkan keuntungan. Memperkenalkan produk lokal ke negara lain dan meningkatkan perekonomian Negara merupakan keuntungan dalam kegiatan ekspor [1]. Salah satu kegiatan ekspor yang memberikan keuntungan bagi, Dimana suatu Negara

memiliki asupan kebutuhan tersendiri terhadap buah-buahan lokal. Di Indonesia, komoditas hortikultura khususnya buah-buahan memiliki potensi besar untuk menambah pundi-pundi devisa negara. Berdasarkan data Badan Pusat Statistik (BPS), ekspor buah pada periode Januari – Desember 2018 naik signifikan hingga mencapai 115,58 persen jika dibanding dengan tahun 2017. Jenis buah yang banyak diekspor antara lain nanas, pisang, dan manggis.

Saat ini Indonesia merupakan peringkat ke-5 dunia sebagai Negara produsen manggis, setelah India, China, Kenya, dan Thailand. Tujuan ekspor nya telah menembus banyak Negara seperti China, Hongkong, Thailand, Vietnam, Singapura, Malaysia, Arab Saudi, Kuwait, Oman, Qatar, Amerika, Australia dan beberapa negara Asean lainnya. Berdasarkan hal tersebut ekspor buah-buahan ke Negara tujuan menjadi hal yang sangat penting karena berpotensi untuk meningkatkan perekonomian Indonesia. Pemetaan terhadap Negara tujuan dengan menggunakan teknik ilmu komputer adalah hal yang dapat dilakukan untuk melihat pengelompokan pada ekspor buah-buahan menurut Negara tujuan. Banyak cabang ilmu komputer yang dapat menyelesaikan permasalahan tersebut. Diantaranya adalah data mining [2].

Penelitian ini menggunakan teknik analisa *clustering*. Ada beberapa penelitian yang menggunakan kelebihan dari *K-Medoids* dalam memecahkan permasalahan seperti yang dilakukan dengan judul Implementasi Metode *K-Medoids Clustering* untuk pengelompokan data potensi kebakaran hutan/lahan berdasarkan persebaran titik panas (*Hotspot*) [3]. Pada penelitian ini data mining *K-Medoids* dapat diterapkan pada pengelompokan data potensi kebakaran hutan/lahan berdasarkan persebaran titik panas (*Hotspot*) dengan hasil-hasil *clustering* menunjukkan bahwa dengan penggunaan 2 *cluster* menghasilkan kelompok data dengan potensi yang dimiliki yaitu *cluster* 1 termasuk dalam potensi tinggi dengan hasil rata-rata *brightness* sebesar 344.47 derajat kelvin dengan rata-rata *confidence* 87.08 persen dan *cluster* 2 masuk dalam potensi sedang dengan hasil rata-rata *brightness* sebesar 318.80 derajat kelvin dengan rata-rata *confidence* sebesar 58,73 persen.

Peneliti *K-Medoids* berikutnya dilakukan dengan judul Optimalisasi *K-Medoids* dalam pengklasteran mahasiswa pelamar beasiswa dengan *Cubic Clustering Criterion*. Pada penelitian ini *K-Medoids* dapat diterapkan dengan memaksimalkan *Cubic Clustering Criterion* pada dalam pengklasteran mahasiswa pelamar beasiswa [4]. Hasil penelitian menunjukkan bahwa *dataset* modifikasi keseluruhan mempunyai keseragaman yang baik. Hal ini dikarenakan semua nilai pada setiap atribut memiliki nilai yang hampir sama. Berdasarkan kelebihan tersebut, solusi yang diberikan pada penelitian ini menggunakan metode *K-Medoids* yang diharapkan hasil penelitian dapat memberikan informasi kepada pemerintah tentang hasil pemetaan berupa *cluster* terhadap Negara tujuan untuk jumlah ekspor buah-buahan.

## **METODE**

Metode penelitian yang dipergunakan dalam penelitian ini yaitu metode kualitatif, karena metode ini bersifat deskriptif dan cenderung analisis yang berupa kata-kata tertulis atau lisan dari perilaku yang diamati. Metode kualitatif lebih sering menggunakan proses mendalam yaitu dengan mempelajari masalah secara kasus per kasus karena metode kualitatif yakin bahwa suatu masalah satu akan berbeda dengan masalah lainnya [5].

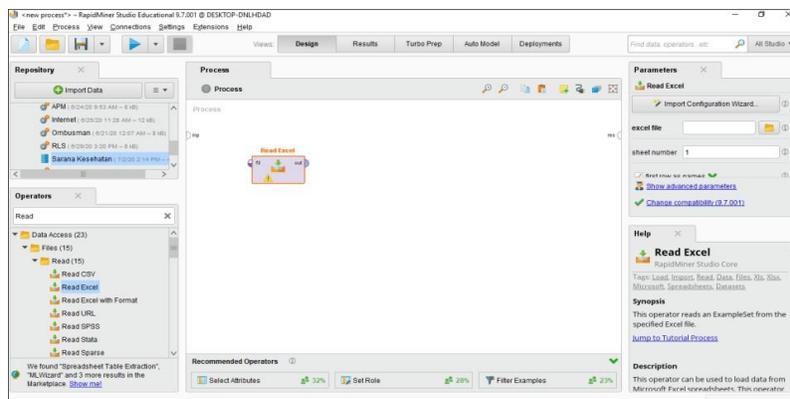
## HASIL DAN PEMBAHASAN

### Halaman Login

Halaman *login* adalah halaman untuk hak atau batasan terhadap *user* untuk masuk kedalam pengolahan sistem. *Username* dan *password* yang sah dimiliki pengguna untuk berhak masuk ke dalam *system* melalui halaman *login*. Gambar di bawah ini adalah tampilan halaman *login*. Bagian dari menu awal pada *tools* Rapidminer seperti tampilan awal dari Rapidminer dan akan dimulai dengan menu *New Process*.

### Input Sistem

Sistem menjelaskan cara memasukan data baru yang akan dieksekusi lebih lanjut, pada hal ini data yang akan di eksekusi berupa data *Excel*. Data yang diproses sebagai *input* diolah menggunakan operator *read excel* yang tentunya data dipersiapkan terlebih dahulu. Data disimpan dengan file ekspor buahan.xls dengan atribut negara asal, jumlah ekspor (netto) dan nilai *Free On Board* (US\$) tahun 2015-2018 (11 *record*). Lakukan *Drag* dan *Drop* pada operator *Read Excel* sehingga seperti gambar berikut.



Gambar 1. *Drag dan Drop* pada operator *Read Excel*

Setelah *Drag* dan *Drop* pada operator *Read Excel*, lakukan *double klik* untuk memasukkan data *excel* (*ekspor* buah-buahan.xls) yang telah disediakan sebelumnya. Ikuti proses *step by step*. Sebenarnya RapidMiner akan memberikan tipe data yang tepat secara otomatis. Namun, jika tipe data yang diberikan RapidMiner tidak cocok, maka tipe data dapat diubah sesuai dengan kebutuhan. Sebelum proses input selesai, pada *format*

column yang harus diperhatikan adalah status atribut negara diubah menjadi *id* karena proses pemetaan berupa *klaster* berdasarkan negara tujuan. Setelah selesai klik *finish*.

Untuk memastikan data yang di *input* benar, hubungkan *output* pada *database* ke *result* seperti Gambar berikut. Lalu klik ikon *Play* seperti yang ditunjukkan pada gambar 2.

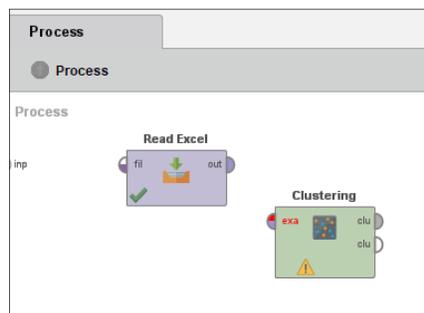
Row No.	NegaraAsal	X1	X2	X3	X4	X5	X6	X7	X8
1	Tiongkok	217120.100	279289.600	215835.900	341863.900	339657.500	564379	397717.300	741351.400
2	AmerikaSerikat	42385.400	106702.800	59091.100	141670.300	46055.600	116226.600	38045.300	110523
3	Thailand	52781.700	72864.900	33895.100	93035.500	100901.100	157770.200	61576.800	112881.900
4	Australia	28965.800	66297.500	60375	85314.600	27053.100	89212.600	27166.300	100704.300
5	Pakistan	20271	23025.700	12784.700	29192.500	46907.500	59007.800	48859.300	69059.100
6	Peru	5189.300	20565.400	7433.700	29166.300	9172.600	36689.500	2730.100	11272.900
7	SelandiaBaru	6614.500	15902.300	17309	20441.100	9366.600	24073.200	6589.500	19880.600
8	Vietnam	10365.200	12733.300	14130.200	19172.900	4846.500	12077	1560	7050.800
9	AfrikaSelatan	9721.900	11494.600	10194.300	16428	14895.100	25168	11711.500	22550.200
10	Mesir	12010.700	10951.600	11891.300	10703.100	20269.500	21585.700	23975.400	27461.600
11	Lainnya	28578.500	46745.300	37488.300	61149	44685.600	85397.700	947610.900	826419.400

Gambar 2. Isi data penelitian pada Rapid Miner

Pada gambar 2 dapat dijelaskan bahwa isi data penelitian dapat kita lihat dengan cara menghubungkan *output* repositori ke *result* seperti yang ditunjukkan pada gambar 5.8 Untuk melihat data lengkap atribut yang digunakan, klik icon *statistic* sehingga tampil data lengkap atribut mulai dari tipe data, *missing*, *statistics*, *filter*, *value* seperti yang ditunjukkan pada gambar berikut:

Untuk membuat *k-medoids* dengan menggunakan RapidMiner, maka dibutuhkan operator *k-medoids* yang terdapat pada *View Operators*. Untuk menggunakannya pilih *Modelling* pada *View Operator*, lalu pilih *Segmentation*, lalu pilih *k-medoids*.

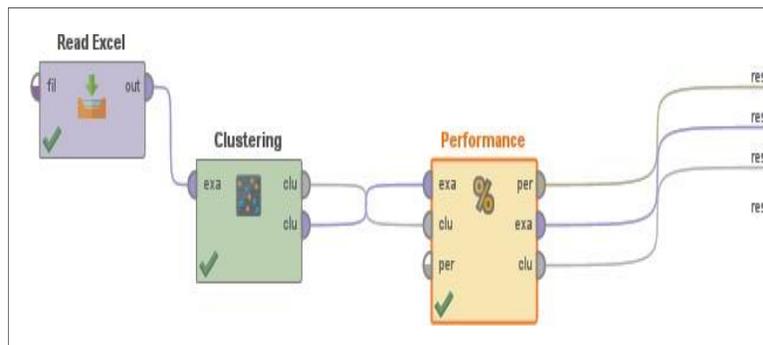
Setelah menemukan operator *k-medoids*, seret (*drag*) operator tersebut lalu letakkan (*drop*) ke dalam *viewProcess*. Kemudian susun posisinya disamping operator *Read Excel*, seperti yang tampak pada gambar 3.



Gambar 3. Posisi Operator *K-Medoid*

Selanjutnya, hubungkan operator *Read Excel* dengan operator *k-medoids* dengan menarik garis dari *Read Excel* ke operator *k-medoids*. Dalam hal ini penulis

menggunakan operator *cluster distance performance* dengan cara melakukan *drag* dan *drop* kedalam *viewProcess*. Kemudian menarik garis lagi dari operator *k-medoids* ke *cluster distance performance* dan dari *cluster distance performance* menarik garis ke *result* di sisi kanan, seperti yang tampak pada Gambar berikut: Operator *k-medoids* berguna untuk melakukan pengelompokan bersama objek yang mirip satu sama lain dan berbeda dengan objek milik klaster lain. Sementara operator *cluster distance performance* berguna untuk evaluasi kinerja metode pengelompokan berbasis *centroid*. Operator ini memberikan daftar nilai kriteria kinerja berdasarkan *centroid* klaster.



Gambar 4. Menghubungkan tabel *ekspor* buah.xls dengan *k-medoids* dan *cluster distance performance*

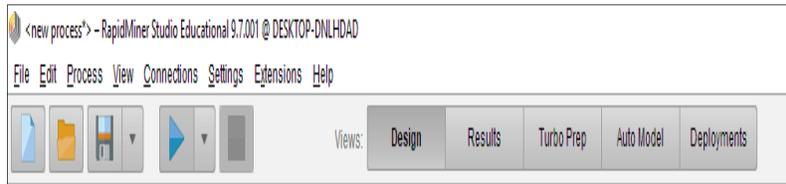
Langkah selanjutnya adalah mengatur parameter sesuai dengan kebutuhan. Setelah menghubungkan operator *Read Excel* dengan operator *k-medoids*, atur parameter *k-medoids* seperti pada gambar berikut.

### Proses Sistem

Pada tahap ini dijelaskan tahapan proses penggunaan *k-medoids* di dalam *RapidMiner*. Seperti yang dijelaskan pada gambar 4, operator *k-medoids* tidak langsung dihubungkan ke *results*. Melainkan dihubungkan ke operator *cluster distance performance* yang berfungsi untuk mengevaluasi kinerja metode pengelompokan berbasis *centroid*. Operator ini memberikan daftar nilai kriteria kinerja berdasarkan *centroid cluster*. Parameter yang digunakan adalah *Davies Bouldin Index (DBI)*. DBI sebagai acuan pengelompokan klaster terbaik dengan melihat nilai minimal dari nilai *k* yang diberikan.

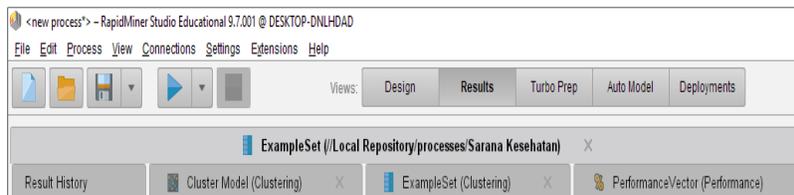
### Output Sistem

Setelah parameter diatur, klik ikon *Run* pada *toolbar*. Tunggu beberapa saat, *compute* membutuhkan waktu untuk menyelesaikan perhitungan seperti yang ditunjukkan pada gambar 5.



Gambar 5. Ikon Run

Setelah beberapa detik maka RapidMiner akan menampilkan hasil keputusan pada *view Result*. Hasil yang ditampilkan RapidMiner dapat dipilih sesuai dengan kebutuhan. Pada penelitian ini diperoleh hasil *Cluster Model*, *ExampleSet* Dan *Performancevector* seperti yang ditunjukkan pada gambar berikut:



Gambar 6. Tampilan pilihan hasil dengan RapidMiner

Pada tab *cluster model (clustering)* dapat ditampilkan beberapa informasi seperti *description*, *folder view*, *graph*, *centroid table* dan *plot*. Berikut penjelasan gambar dari masing masing informasi yang ditampilkan oleh RapidMiner.

Attribute	cluster_0	cluster_1
X1	12010.700	28578.500
X2	10951.600	46745.300
X3	11891.300	37488.300
X4	10703.100	61149
X5	20269.500	44885.600
X6	21565.700	85397.700
X7	23975.400	947610.900
X8	27461.600	826419.400

Gambar 7. Hasil *centroid* akhir

Pada gambar 7. adalah hasil *centroid* akhir dari perhitungan *k-medoids* menggunakan *software* RapidMiner. Hasil *centroid* akhir ini sama dengan hasil dari perhitungan manual. Pada tab *exampleset (clustering)* dapat ditampilkan beberapa informasi seperti *data*, *statistics* dan *visualizations*. Berikut penjelasan gambar dari masing masing informasi yang ditampilkan oleh RapidMiner.

Row No.	NegaraAsal	label	X1	X2	X3	X4	X5	X6	X7	X8
1	Tiongkok	cluster_1	217120.100	279289.600	215835.900	341863.900	339657.500	564379	397717.300	741351.400
2	AmerikaSerikat	cluster_0	42385.400	106702.800	59091.100	141670.300	46055.600	116226.600	38045.300	110523
3	Thailand	cluster_0	52781.700	72664.900	33995.100	93035.500	100901.100	157770.200	61576.800	112881.900
4	Australia	cluster_0	29965.800	66297.500	60375	85314.600	27053.100	89212.600	27166.300	100704.300
5	Pakistan	cluster_0	20271	23025.700	12784.700	29192.500	46907.500	59007.800	48859.300	69059.100
6	Peru	cluster_0	5189.300	20565.400	7433.700	29166.300	9172.600	36689.500	2730.100	11272.900
7	SelandiaBaru	cluster_0	6614.500	15902.300	17309	20441.100	9366.600	24073.200	6589.500	19880.600
8	Vietnam	cluster_0	10365.200	12733.300	14130.200	19172.900	4846.500	12077	1560	7050.800
9	AfrikaSelatan	cluster_0	9721.900	11494.600	10194.300	16428	14895.100	25168	11711.500	22550.200
10	Mesir	cluster_0	12010.700	10951.600	11891.300	10703.100	20269.500	21565.700	23975.400	27461.600
11	Lainnya	cluster_1	28578.500	46745.300	37488.300	61149	44865.600	85397.700	947610.900	826419.400

Gambar 8. Hasil lengkap kluster dengan RapidMiner

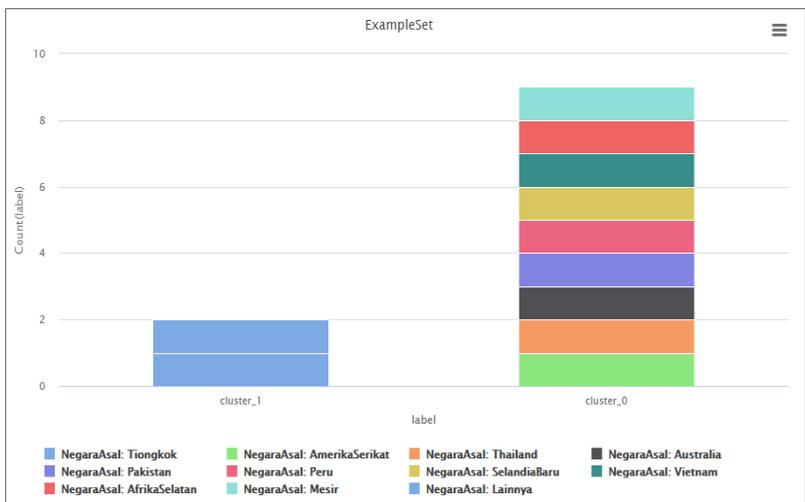
Gambar 9. Proses *export* hasil kluster dengan RapidMiner ke *Excel*

Hasil pengelompokan dengan *k-medoids* dapat dilakukan dengan menggunakan RapidMiner. Hasil dari data pengelompokan dapat di *export* ke *Microsoft Excel* seperti yang ditunjukkan pada tabel berikut ini.

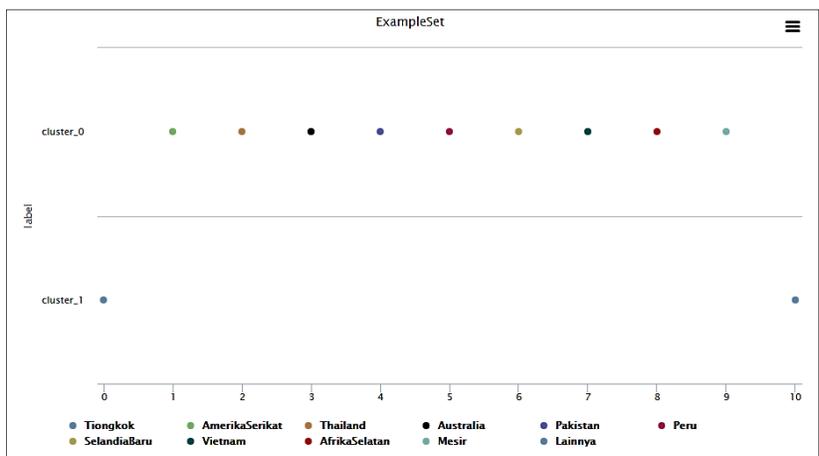
Name	Type	Missing	Statistics	Filter (10 / 10 attributes)
NegaraAsal	Polynomial	0	Least Vietnam (1)	Most AfrikaSelatan (1)
label	Nominal	0	Least cluster_1 (2)	Most cluster_0 (9)
X1	Real	0	Min 5189.300	Max 217120.100
X2	Real	0	Min 10951.600	Max 279289.600
X3	Real	0	Min 7433.700	Max 215835.900
X4	Real	0	Min 10703.100	Max 341863.900
X5	Real	0	Min 4846.500	Max 339657.500

Gambar 10. Detail dari atribut yang digunakan

Pada gambar 10 dijelaskan detail dari atribut yang digunakan mulai dari nama, *type data*, *missing*, *statistics*, batasan dari *type data* yang digunakan, nilai minimum dan maksimum dari *record data*. Untuk melihat detail data cukup dengan klik salah satu atribut.



Gambar 11. Visualisasi grafik dengan *Bar*

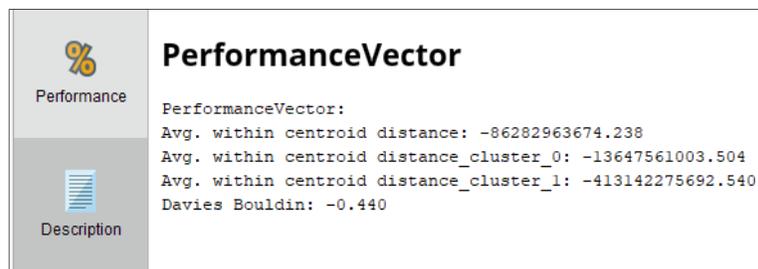


Gambar 12. Visualisasi grafik dengan *scatter*

Gambar 11 dan 12 merupakan pengelompokan ekspor buah-buahan menurut negara tujuan yang terdiri dari 2 kluster yakni kluster tinggi ( $C1 = cluster_1$ ), dan kluster rendah ( $C1 = cluster_0$ ). Pada tab *performavector* (*performance*) dapat ditampilkan beberapa informasi seperti *performance* dan *description*. Berikut penjelasan gambar dari masing masing informasi yang ditampilkan oleh RapidMiner.

Gambar 13. Hasil *Davies Bouldin Index* (DBI)

Berdasarkan nilai *Davies-Bouldin Index* (DBI) diatas, menunjukkan bahwa  $k=2$  menjadi *klaster* terbaik. Nilai DBI = 0.440 merupakan klastering yang paling optimal dengan nilai DBI minimal.



Gambar 14. Hasil Lengkap PerformanceVector

## SIMPULAN

Berdasarkan hasil penelitian dapat disimpulkan bahwa pengelompokan ekspor buah-buahan menurut negara tujuan dapat diterapkan dengan metode *k-medoids* dimana terdapat 2 label klaster yang digunakan yakni klaster tinggi ( $C1= cluster_1$ ) dan klaster rendah ( $C2= cluster_0$ ). Hasil *klaster* menyebutkan bahwa 2negara berada di *klaster* tinggi ( $C1= cluster_1$ ), 9 negara berada di klaster rendah ( $C2= cluster_0$ ). (a) *Klaster* rendah ( $C2=cluster_0$ ) yakni Amerika Serikat, Thailand, Australia, Pakistan, Peru, Selandia Baru, Vietnam, Afrika Selatan, Mesir., (b) *Klaster*tinggi ( $C1=luster_1$ ) yakni Tiongkok dan Lainnya.

## DAFTAR PUSTAKA

- [1] D. F. Pramesti, Lahan, M. Tanzil Furqon, and C. Dewi, "Implementasi Metode K-Medoids Clustering Untuk Pengelompokan Data," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 1, no. 9, pp. 723–732, 2017, doi: 10.1109/EUMC.2008.4751704.
- [2] Z. Mustofa and I. S. Suasana, "Algoritma Clustering K-Medoids Pada E-Government Bidang Information And Communication," *J. Teknol. dan Komun.*, vol. 9, pp. 1–10, 2018.

- [3] A. Y. Rofiqi, “Clustering Berita Olahraga Berbahasa Indonesia Menggunakan Metode K-Medoid Bersyarat,” *J. Simantec*, vol. 6, no. 1, pp. 25–32, 2017.
- [4] K. Kawano, Y. Umemura, and Y. Kano, “ Field Assessment and Inheritance of Cassava Resistance to Superelongation Disease 1 ,” *Crop Sci.*, vol. 23, no. 2, pp. 201–205, 1983, doi: 10.2135/cropsci1983.0011183x002300020002x.
- [5] W. A. Triyanto, “Algoritma K-Medoids Untuk Penentuan Strategi Pemasaran Produk,” *Simetris J. Tek. Mesin, Elektro dan Ilmu Komput.*, vol. 6, no. 1, p. 183, 2015, doi: 10.24176/simet.v6i1.254.