

Penerapan Teknik *Clustering* untuk Mendukung Diferensiasi Strategi Pengelolaan Bandar Udara

Fadhlan Nabil Nabassa, Fahmi Arif, ST., MT., Ph.D.

Institut Teknologi Nasional Bandung
E-mail : fadhlan437@mhs.itenas.ac.id

Received DD MM YYYY | Revised DD MM YYYY | Accepted DD MM YYYY

ABSTRAK

Salah satu pengelola bandar udara di Indonesia yang mengelola sebanyak 20 bandar udara memerlukan strategi pengelolaan bandar udara yang spesifik bagi setiap bandar udara. Penentuan strategi pengelolaan bandar udara secara individu sulit dilakukan mengingat jumlah bandara yang cukup banyak dengan perbedaan karakteristik yang cukup signifikan. Salah satu alternatif untuk menyelesaikan masalah tersebut adalah pengelompokan yang akan menjadi acuan bagi pengelola bandara dalam menentukan strategi pengelolaan bandar udara. Pengelompokan dilakukan dengan mempertimbangkan performansi bandar udara. Hasil clustering menunjukkan metode k-means dan hierarchical clustering memberikan hasil yang optimum dengan jumlah klaster sebanyak tiga klaster dan nilai silhouette score 0,678.

Kata kunci: bandara, crisp-dm, pengelompokan, k-means, hierarchical clustering

ABSTRACT

One airport manager in Indonesia who manages as many as 20 airports requires a specific airport management strategy for each airport. Determining an individual airport management strategy is difficult given the large number of airports with significant differences in characteristics. One alternative to solve this problem is grouping which will become a reference for airport managers in determining airport management strategies. The grouping is done by considering the performance of the airport. The clustering results show that the k-means and hierarchical clustering methods provide optimum results with the number of clusters being three clusters and the silhouette score of 0.678

Keywords: airport, crisp-sm, clustering, k-means, hierarchical clustering

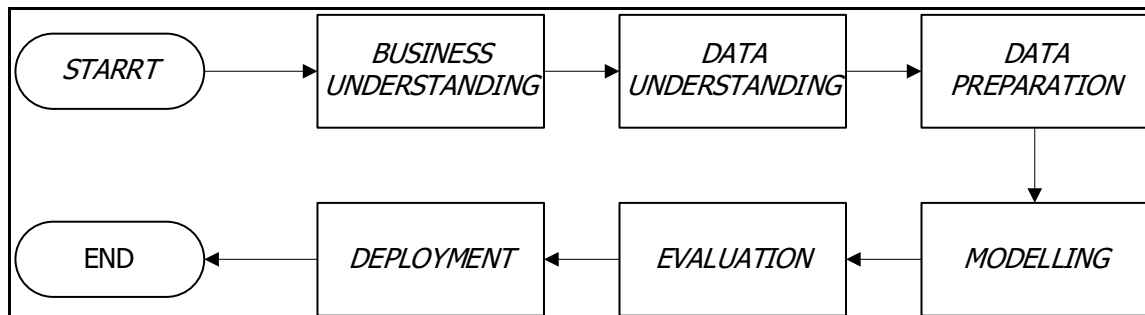
1. PENDAHULUAN

Bandar udara menurut Undang – Undang Republik Indonesia No. 1 Tahun 2009 Tentang Penerbangan adalah “kawasan di daratan dan/atau perairan dengan batas-batas tertentu yang digunakan sebagai tempat pesawat udara mendarat dan lepas landas, naik turun penumpang, bongkar muat barang, dan tempat perpindahan intra dan antarmoda transportasi, yang dilengkapi dengan fasilitas keselamatan dan keamanan penerbangan, serta fasilitas pokok dan fasilitas penunjang lainnya”. Menurut Peraturan Menteri Perhubungan No. 69 Tahun 2013 Tentang Tatanan Kebandarudaraan Nasional terdapat dua fungsi bandar udara, yaitu fungsi pemerintahan sebagai tempat unit kerja pemerintah dalam melaksanakan fungsi dan tugas kepada masyarakat sesuai peraturan yang berlaku serta sebagai tempat untuk melakukan perusahaan. Tujuan dari pengelolaan bandar udara adalah untuk memajukan bandar udara yang bersangkutan dengan cara memberikan pelayanan yang maksimal kepada pelanggan serta untuk meningkatkan pertumbuhan ekonomi masyarakat di sekitar bandar udara **(Setiani, 2015)**.

Salah satu pengelola bandar udara di Indonesia mengelola sebanyak 20 bandar udara **(Puspa, 2021)**. Setiap bandar udara yang dikelola memiliki karakteristik serta performansinya masing – masing. Oleh karena itu diperlukan diferensiasi strategi pengelolaan bandar udara yang sesuai dengan karakteristik masing – masing. Banyaknya bandar udara yang dikelola serta memiliki karakteristik dan performansinya masing – masing membuat penentuan strategi pengelolaan yang spesifik bagi setiap bandar udara memiliki kompleksitas yang tinggi. Faktor – faktor yang mempengaruhi penentuan strategi pengelolaan bandar udara adalah faktor keselamatan, kapasitas infrastruktur, jumlah sumber daya manusia yang tersedia, jumlah jadwal keberangkatan, dan adaptasi terhadap perubahan yang terjadi **(Alodhaibi et al., 2017)**.

Salah satu upaya yang dapat dilakukan untuk memudahkan penentuan strategi pengelompokan bandar udara adalah mengelompokkan bandar udara berdasarkan kemiripan karakteristiknya. Penggunaan metode pengelompokan tersebut diharapkan dapat membantu dalam menentukan strategi pengelolaan bandar udara yang spesifik untuk setiap bandar udara. Selain itu jumlah bandar udara yang dikelola dapat mengalami perubahan sesuai dengan ketetapan Pemerintah sehingga pengelompokan ini juga diharapkan dapat memudahkan dalam menentukan strategi pengelolaan bandar udara yang baru atau akan dikelola. Bandar udara GGK dan KNO memiliki karakteristik yang berbeda dengan bandar udara lainnya sehingga kedua bandar udara tersebut tidak termasuk kedalam bandar udara yang dikelompokkan. Tingkat akurasi yang tepat diperlukan agar dapat menentukan strategi pengelolaan yang tepat bagi setiap bandar udara, oleh karena itu penentuan variabel yang digunakan dalam pengelompokan serta penentuan hasil metode pengelompokan yang optimum perlu dikaji agar memberikan hasil pengelompokan yang akurat. Penelitian ini memiliki tujuan untuk menghasilkan pengelompokan bandar udara berdasarkan kemiripan karakteristiknya dengan menggunakan teknik *partitional clustering* dan *hierarchical clustering* untuk mendukung diferensiasi strategi pengelolaan bandar udara.

2. METODOLOGI



Gambar 1. Diagram Alir Metodologi Penelitian

Metodologi penelitian yang digunakan mengadopsi metode *Cross Industry Standard Process for Data Mining* (CRISP-DM). CRISP-Dm merupakan standar model proses untuk penerapan *data mining* yang terdiri dari enam fase, yaitu *business understanding*, *data understanding*, *data preparation*, *modelling*, *evaluation*, dan *deployment* (Schröer et al., 2021). Tahapan pada proses CRISP-DM dapat dilihat pada Gambar 1.

Tahapan pertama yang dilakukan adalah *business understanding*. Pada tahapan ini dilakukan proses analisis untuk memahami mengenai proses bisnis bandar udara, mekanisme pengelolaan bandar udara, dasar dalam melakukan pengelompokan, serta mengidentifikasi variabel performansi bandar udara. Berdasarkan hasil analisis diketahui bahwa kegiatan usaha pengelola bandar udara tidak hanya berkaitan dengan jasa kebandarudaraan tetapi termasuk juga pada kegiatan usaha penunjang pelayanan untuk penumpang. Bandar udara perlu dikelola secara efektif, efisien dan andal, hal ini disebabkan karena ketiga hal tersebut mempengaruhi pelayanan terhadap penumpang dan sebagai alat ukur kinerja pengelolaan bandar udara. Perencanaan strategi pengelolaan bandar udara yang spesifik yang untuk setiap bandar udara memiliki kompleksitas yang tinggi sehingga dilakukan proses pengelompokan berdasarkan kemiripan karakteristiknya yang nantinya akan menjadi acuan dalam menentukan strategi pengelolaan bandar udara. Variabel yang digunakan pada pengelompokan adalah *revenue aero* dan *non – aero* serta pergerakan penumpang dan pesawat. Keempat variabel tersebut dipilih karena bersifat dinamis sehingga berubah setiap tahunnya. Hal ini dilakukan agar dapat melihat perkembangan pengelompokan setiap tahunnya.

Tahapan kedua adalah *data understanding*. Pada tahapan ini dilakukan proses identifikasi atribut dari performansi bandar udara, identifikasi kebutuhan data, pengumpulan data, dan pendeskripsian data yang diperoleh. Terdapat enam atribut yang menjadi indikator performansi bandar udara, yaitu *core*, *safety* dan *security*, *service quality*, *productivity*, *financial*, dan *environmental* (Hazel, 2012). Atribut yang digunakan pada pengelompokan adalah *core* dan *financial*. Atribut *core* dipilih karena digunakan untuk melihat pengukuran hal – hal yang *fundamental* bagi aktivitas bandar udara seperti pergerakan penumpang dan pesawat sedangkan atribut *financial* dipilih karena digunakan untuk melihat performansi finansial bandar udara, seperti *revenue aero* dan *non – aero*. Data yang digunakan pada penelitian ini adalah data pada tahun 2018 hingga 2020. Data yang digunakan berasal dari data sekunder pihak pengelola bandar udara. Berikut ini adalah pendeskripsian atribut yang digunakan pada penelitian ini yang dapat dilihat pada Tabel 1.

Tabel 1. Deskripsi Data yang Digunakan

Data	Satuan	Type Data	Penjelasan Data
<i>Revenue Aero</i>	Rupiah per tahun	<i>Numeric</i>	Data pendapatan bandar udara yang berasal dari kegiatan jasa penerbangan selama satu tahun
<i>Revenue Non - Aero</i>	Rupiah per tahun	<i>Numeric</i>	Data pendapatan bandar udara yang berasal dari kegiatan jasa penerbangan selama satu tahun
Pergerakan Penumpang	Orang per tahun	<i>Numeric</i>	Data pergerakan penumpang di bandar udara selama satu tahun
Pergerakan Pesawat	Pesawat per tahun	<i>Numeric</i>	Data pergerakan pesawat di bandar udara selama satu tahun

Tahapan ketiga adalah *data preparation*. Tahapan ini terdiri dari tiga proses, yaitu *data cleaning*, *data selection*, dan *data transformation*. Tahapan *data cleaning* merupakan tahapan untuk menghilangkan data tidak lengkap, data *error*, *outliers*, dan data tidak konsisten (Alasadi & Bhaya, 2017). Pada penelitian ini *data cleaning* digunakan untuk menghilangkan data tidak memiliki nilai atau *missing value* pada tahun 2018 dan 2019 akibat adanya penambahan bandar udara yang dikelola pada tahun 2019 dan 2020. *Data selection* dilakukan untuk memilih data yang digunakan pada proses selanjutnya (*modelling*). Data yang digunakan pada tahapan *modelling* adalah data nama atau kode bandar udara. Selain itu *data selection* juga digunakan untuk memilih data pertahun yang digunakan pada pengelompokan. Hal ini karena data yang didapatkan merupakan data gabungan mulai dari 2018 hingga 2020 untuk setiap atributnya. *Data transformation* digunakan untuk mengubah bentuk data kedalam kedalam *array* dan merubah rentang data dengan menggunakan *minmaxscaler*. Hal ini dilakukan karena data memiliki rentang yang berbeda (Nasution et al., 2019) dan untuk mempercepat perhitungan algoritma (Thara et al., 2019). Proses *data preparation* menggunakan Bahasa pemrograman *python* pada *jupyter notebook*.

Tahapan keempat adalah *modelling*. Pada tahapan ini dilakukan proses *clustering* dengan menggunakan metode *k-means* dan *hierarchical clustering* dengan menggunakan bahasa pemrograman *python* pada *jupyter notebook*. Sama halnya dengan *data preparation*, seluruh proses pada tahapan *modelling* dilakukan dengan menggunakan bahasa pemrograman *python* dan dituliskan pada *jupyter notebook*. Dalam proses *input data*, proses *clustering* sampai dengan menampilkan hasil *clustering* digunakan *python library* seperti *Numpy*, *Pandas*, *Matplotlib*, *Scipy*, serta *Scikit-learn*. Hal ini bertujuan untuk memudahkan dalam melakukan penulisan *syntax* pemrograman, perhitungan yang kompleks, hingga menampilkan data. Tahap awal yang dilakukan adalah menentukan jumlah kluster yang optimal untuk metode *k-means* dengan metode *elbow*. Metode *elbow* merupakan metode yang digunakan untuk menentukan jumlah kluster (Orisa & Faisol, 2021). Metode ini membandingkan nilai WCSS (*Within Clusters Sum of Square*) dengan jumlah kluster yang terbentuk. WCSS (*Within Clusters Sum of Square*) merupakan jarak antara data dengan *centroid*nya dalam satu kluster (Orisa

& Faisol, 2021). Perhitungan nilai WCSS menggunakan rumus sebagai berikut (**Orisa & Faisol, 2021**):

$$WCSS = \sum_{k=1}^k \sum_{xi \in ck} (xi - \mu_k)^2 \quad (1)$$

Ket:

k = Klaster

xi = Data ke - i pada klaster k

μ_k = Nilai centeroid klaster k

Setelah diketahui jumlah klaster yang optimum untuk metode *k-means* maka dilakukan proses pengelompokan yang dilakukan oleh algoritma *clustering k-means* yang terdapat pada *Scikit-learn*. Hasil yang didapatkan pada akan ditampilkan dalam bentuk tabel. Setelah pengelompokan metode *k-means* maka selanjutnya adalah penentuan jumlah klaster yang optimum untuk metode *hierarchical clustering* dengan menggunakan metode *dendrogram*. *Dendrogram* adalah representasi hirarki dari objek yang digunakan untuk menentukan hasil *hierarchical clustering* (**Kansal et al., 2018**). Hasil dari *dendrogram* ditampilkan dalam bentuk grafik. Setelah jumlah klaster yang optimum untuk metode *hierarchical clustering* maka dilakukan proses pengelompokan yang dilakukan oleh algoritma *clustering hierarchical clustering* yang terdapat pada *Scikit-learn*. Setelah setiap tahap pengelompokan selesai dilakukan maka nilai *silhouette score* akan dihitung dengan algoritma menghitung nilai *silhouette score* yang terdapat pada *Scikit-learn*. Menurut Jun (2011), STHDA (2019), dan Rousseeuw (2008) *silhouette score* merupakan metode pengukuran jarak rata – rata antar klaster (**Mangortey et al., 2020**). Perhitungan nilai *silhouette score* dapat dilihat pada perhitungan berikut (**Kansal et al., 2018**):

$$Silhouette\ score = \frac{b-a}{\max(b,a)} \dots \dots \dots (3.7)$$

Ket:

a = Jarak rata – rata *centroid* klaster dengan data yang ada di dalamnya

b = Jarak rata – rata data klaster satu dengan data pada klaster terdekat

Tahapan kelima adalah *evaluation*. Pada tahapan ini dilakukan proses analisis performansi klaster dengan membandingkan nilai *silhouette score*nya serta penentuan hasil *clustering* yang optimum. Hal ini bertujuan untuk mengetahui hasil *clustering* yang optimum. Nilai *silhouette score* digunakan untuk mengukur tingkat akurasi dari hasil pengelompokan dengan rentang nilai yang digunakan -1 sampai 1, nilai semakin mendekati 1 maka hasil pengelompokan semakin baik (**Mangortey et al., 2020**). Penentuan hasil *clustering* yang optimum dilakukan dengan melihat nilai *silhouette score* tertinggi. Hasil *clustering* yang optimum akan dilakukan analisis untuk mengetahui karakteristiknya pada tahapan selanjutnya (*modelling*).

Tahapan keenam adalah *deployment*. Pada tahapan ini dilakukan proses analisis karakteristik klaster yang optimum serta kesimpulan. Analisis dilakukan sehingga dapat ditemukan karakteristik dari bandar udara untuk masing – masing klaster. Analisis dilakukan dengan menggunakan nilai *centroid*. Dilakukan proses pengkategorisasi setiap variabel menggunakan nilai *centroid* menjadi lima kategori. Menurut Sudijono (2012) pada jurnal **Hanief & Purnomo (2019)** menjelaskan mengenai perhitungan pembagian rentang nilai

lima kategori yang dapat dilihat pada Tabel 2. Kesimpulan berisikan mengenai hasil yang didapatkan pada penelitian ini serta saran untuk penelitian selanjutnya.

Tabel 1 Rumus Perhitungan Rentang Lima Kategori

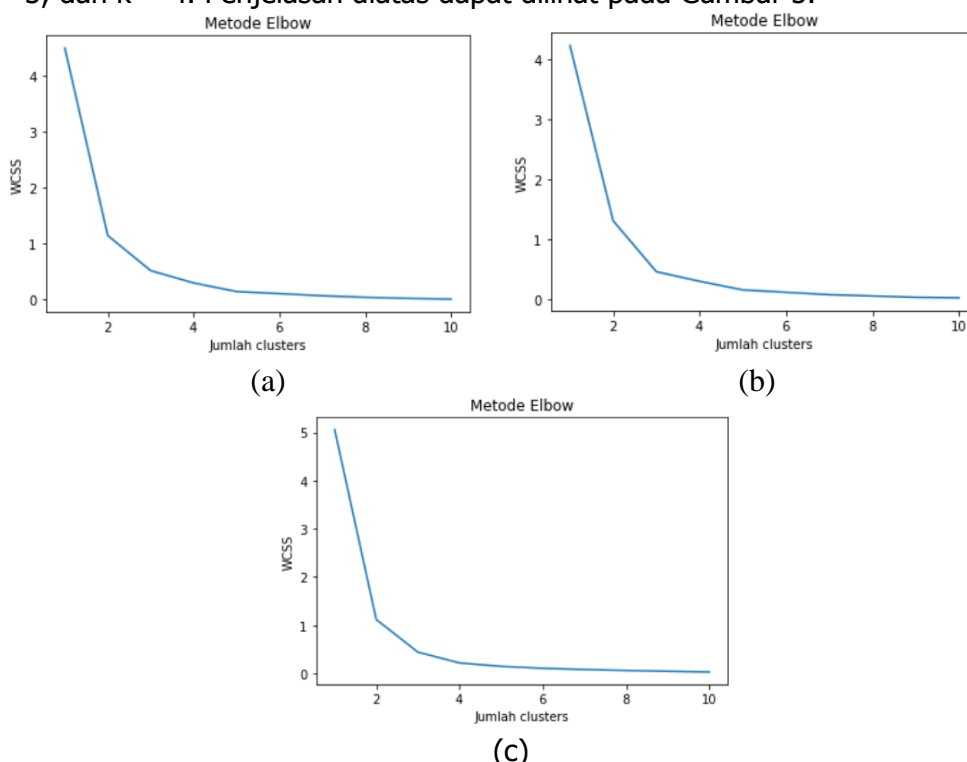
No	Interval	Kategori
1	$M + 1,5 S < X$	Sangat Tinggi
2	$M + 0,5 S < X < M + 1,5 S$	Tinggi
3	$M - 0,5 S < X < M + 0,5 S$	Sedang
4	$M - 1,5 S < X < M - 0,5 S$	Rendah
5	$X < M - 1,5 S$	Sangat Rendah

M = Nilai rata - rata (mean)
S = Standar Deviasi

3. HASIL DAN PEMBAHASAN

Keempat variabel yang digunakan merupakan data yang bersifat dinamis. Hal ini dikarenakan performansi bandar udara yang berubah setiap tahunnya. Perubahan performansi tersebut dipengaruhi oleh kegiatan penerbangan, kegiatan komersial, serta peran pemerintah dan pengelola bandar udara. Pada tahun awal tahun 2020 Indonesia terkena wabah virus COVID-19 yang membuat aktivitas penerbangan mengalami penundaan. Hal ini menyebabkan pendapatan dari PT. Angkasa Pura II mengalami penurunan.

Penggunaan metode *elbow* adalah untuk melihat jumlah kluster yang optimum untuk metode *clustering k-means*. Berdasarkan hasil perhitungan WCSS diketahui terbentuk siku pada $k = 2$, $k = 3$, dan $k = 5$ untuk tahun 2018 dan 2019 sedangkan tahun 2020 terbentuk siku pada $k = 2$, $k = 3$, dan $k = 4$. Penjelasan diatas dapat dilihat pada Gambar 5.



Gambar 1 (a) Metode *Elbow* 2018. (b) Metode *Elbow* Tahun 2019, dan (c) Metode *Elbow* Tahun 2020

Pada Gambar 5 diketahui bahwa terjadi penurunan nilai WCSS untuk setiap tahunnya secara signifikan mulai dari $k = 2$ sampai $k = 3$ yang membentuk sebuah siku. Pada $k = 4$ dan

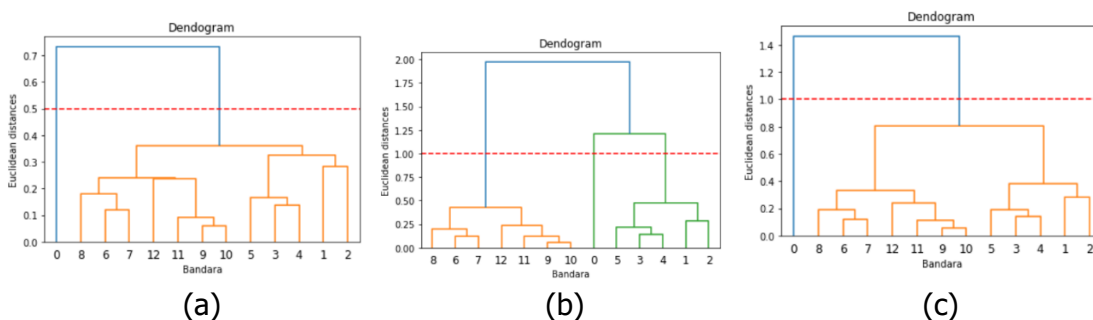
seterusnya penurunan yang nilai yang terjadi sudah tidak signifikan. Hal tersebut menunjukkan bahwa jumlah kluster yang optimum adalah tiga kluster. Pembentukan siku pada kurva menunjukkan jumlah kluster yang optimum. Setelah jumlah k diketahui maka selanjutnya dilakukan proses *clustering k-means* dengan menggunakan algoritma pengelompokan *k-means* pada *Scikit-learn*. Hasil yang didapatkan akan ditampilkan dalam bentuk tabel. Penjelasan mengenai tabel dapat dilihat pada Tabel 6.

Tabel 6 Hasil *Clustering* Metode *K-Means*

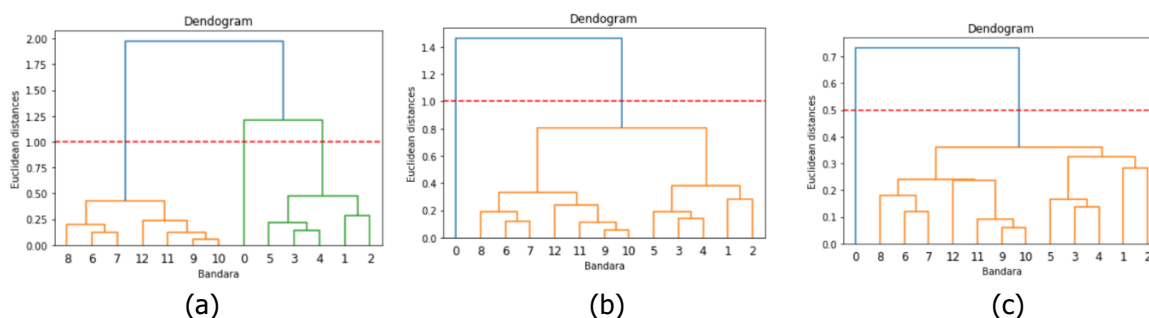
<i>K-Means</i>			
Tahun	Kluster 0	Kluster 1	Kluster 2
2018	HLP, PLM (2)	PGK, DJB, BTJ, TNJ, DTB, BWX (6)	PKU, PDG, BDO, PNK (4)
2019	PGK, DJB, BTJ, TNJ, DTB, BWX, PKY (7)	HLP (1)	PLM, PKU, PDG, BDO, PNK (5)
2020	PLM, PKU, PDG, PNK (4)	BDO, PGK, DJB, BTJ, TNJ, DTB, BWX, PKY, TJQ, BKS, TKG (11)	HLP (1)

Berdasarkan Tabel 6 diketahui bahwa terdapat perbedaan anggota pada setiap kluster walaupun memiliki jumlah kluster yang sama. Setelah kelompok bandar udara terbentuk maka dilakukan proses perhitungan nilai *silhouette score*. *Silhouette score* merupakan cara untuk mengetahui nilai pengelompokan yang terbentuk (**Kansal et al., 2018**). Nilai *silhouette score* tahun 2018 dan 2019 adalah 0,543 dan tahun 2020 adalah 0,678.

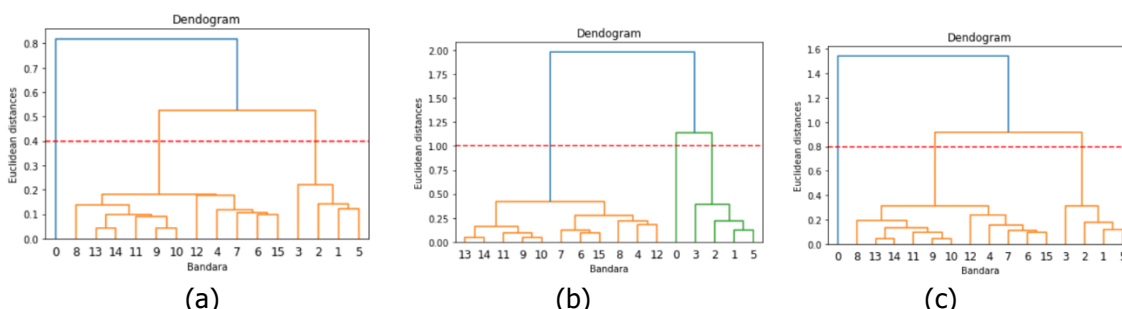
Penentuan jumlah kluster yang optimum untuk metode *hierarchical clustering* dengan menggunakan *dendrogram*. Pada *dendrogram* terdapat sumbu x dan sumbu y dimana sumbu x merupakan bandar udara dan sumbu y merupakan *euclidean distance* yang digunakan untuk menghitung jarak data terhadap masing – masing *centroid* (**Dinata et al., 2020**). Penggabungan antar bandar udara dilakukan berdasarkan perhitungan jarak pada metode *hierarchical clustering* yang digunakan, yaitu *single*, *complete*, dan *average*. Metode *single* menghitung jarak terdekat dua data setiap kluster, metode *complete* menghitung jarak terjauh dua data setiap kluster, dan metode *average* menghitung jarak rata – rata setiap data pada satu kluster dengan setiap data pada kluster lainnya (**Mangortey et al., 2020**). Penentuan jumlah kluster didapat dari jumlah garis vertikal terpanjang yang memotong garis acuan horizontal (**Kansal et al., 2018**). Penempatan garis acuan dilakukan secara acak tetapi tetap memotong garis vertikal terpanjang pada *dendrogram* (**Kansal et al., 2018**). Garis acuan pada *dendrogram* ditampilkan dalam garis putus – putus berwarna merah. Hasil *dendrogram* tahun 2018 dapat dilihat pada Gambar 6 tahun 2019 Gambar 7, dan tahun 2020 Gambar 8.



Gambar 2 (a) *Dendrogram Hierarchical Clustering "single"*, (b) *Dendrogram Hierarchical Clustering "complete"*, dan (c) *Dendrogram Hierarchical Clustering "average"*



Gambar 7 (a) Dendrogram Hierarchical Clustering "single", (b) Dendrogram Hierarchical Clustering "complete", dan (c) Dendrogram Hierarchical Clustering "average"



Gambar 8 (a) Dendrogram Hierarchical Clustering "single", (b) Dendrogram Hierarchical Clustering "complete", dan (c) Dendrogram Hierarchical Clustering "average"

Pada Gambar 6 diketahui bahwa jumlah kluster yang optimum tahun 2018 adalah tiga dan empat kluster. Pada Gambar 7 diketahui jumlah kluster yang optimum tahun 2019 adalah dua dan tiga kluster. Pada Gambar 8 diketahui bahwa jumlah kluster yang optimum tahun 2020 adalah tiga kluster. Setelah jumlah kluster optimum diketahui dilakukan proses pengelompokan metode *hierarchical clustering*. Hasil pengelompokan dapat dilihat pada Tabel 7, Tabel 8, dan Tabel 9.

Tabel 7 Hasil Clustering Metode Hierarchical Clustering "Single"

Hierarchical Clustering "Single"				
Tahun	Klaster 0	Klaster 1	Klaster 2	Klaster 3
2018	PGK, DJB, BTJ, TNJ, DTB, BWX (6)	PLM (1)	PKU, PDG, BDO, PNK (4)	HLP (1)
2019	PLM, PKU, PDG, BDO, PNK, PGK, DJB, BTJ, TNJ, DTB, BWX, PKY (12)	HLP (1)		
2020	PLM, PKU, PDG, PNK (4)	HLP (1)	BDO, PGK, DJB, BTJ, TNJ, DTB, BWX, PKY, TJQ, BKS, TKG (11)	

Tabel 8 Hasil Clustering Metode Hierarchical Clustering "Complete"

Hierarchical Clustering "Complete"			
Tahun	Klaster 0	Klaster 1	Klaster 2
2018	HLP, PLM (2)	PGK, DJB, BTJ, TNJ, DTB, BWX (6)	PKU, PDG, BDO, PNK (4)
2019	PLM, PKU, PDG, BDO, PNK (5)	PGK, DJB, BTJ, TNJ, DTB, BWX, PKY (7)	HLP (1)
2020	BDO, PGK, DJB, BTJ, TNJ, DTB, BWX, PKY, TJQ, BKS, TKG (11)	PLM, PKU, PDG, PNK (4)	HLP (1)

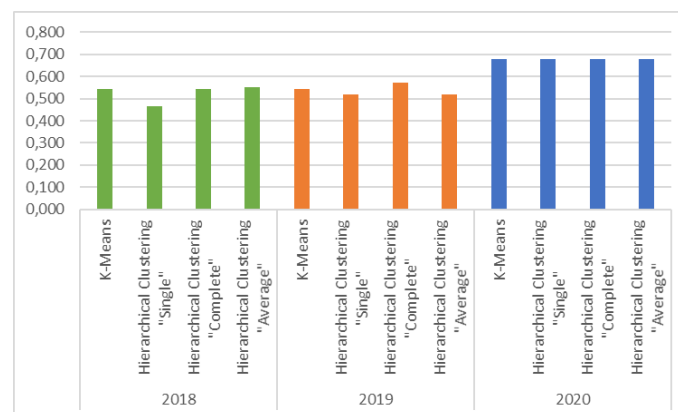
Tabel 9 Hasil *Clustering* Metode *Hierarchical Clustering* "Average"

<i>Hierarchical Clustering</i> "Average"			
Tahun	Klaster 0	Klaster 1	Klaster 2
2018	PLM, PKU, PDG, BDO, PNK (5)	PGK, DJB, BTJ, TNJ, DTB, BWX (6)	HLP (1)
2019	PLM, PKU, PDG, BDO, PNK, PGK, DJB, BTJ, TNJ, DTB, BWX, PKY (12)	HLP (1)	
2020	PLM, PKU, PDG, PNK (4)	HLP (1)	BDO, PGK, DJB, BTJ, TNJ, DTB, BWX, PKY, TJQ, BKS, TKG (11)

Tabel 4.3, Tabel 4.4, dan Tabel 4.5 menunjukkan bahwa hasil yang didapatkan untuk setiap metode pada tahunnya sama memiliki anggota klaster yang berbeda. Selanjutnya dilakukan proses perhitungan nilai *silhouette score*. Berikut adalah nilai *silhouette score* metode *hierarchical clustering* yang dapat dilihat pada Tabel 10 dan nilai *silhouette score* seluruh metode yang dapat dilihat pada Gambar 9.

Tabel 10 Nilai *Silhouette Score* Metode *Hierarchical Clustering*

Tahun	Metode	<i>Silhouette Score</i>
2018	<i>Hierarchical Clustering</i> "Single"	0,466
	<i>Hierarchical Clustering</i> "Complete"	0,543
	<i>Hierarchical Clustering</i> "Average"	0,551
2019	<i>Hierarchical Clustering</i> "Single"	0,519
	<i>Hierarchical Clustering</i> "Complete"	0,572
	<i>Hierarchical Clustering</i> "Average"	0,519
2020	<i>Hierarchical Clustering</i> "Single"	0,678
	<i>Hierarchical Clustering</i> "Complete"	0,678
	<i>Hierarchical Clustering</i> "Average"	0,678



Gambar 9 Nilai *Silhouette Score* Seluruh Metode

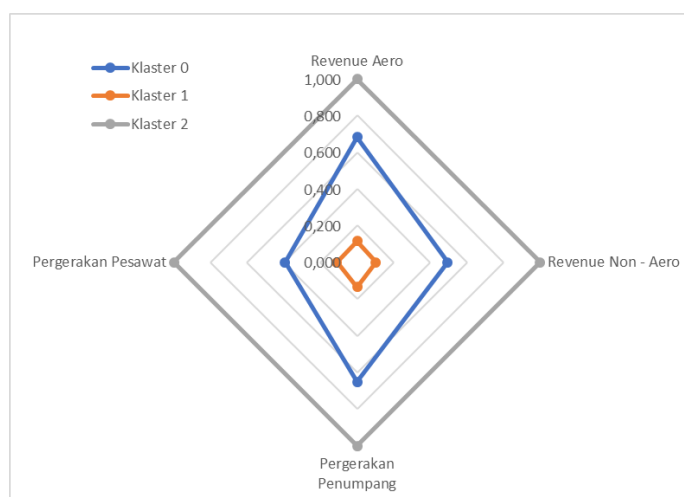
Pada Tabel 10 diketahui nilai *silhouette score* tertinggi dihasilkan oleh metode *hierarchical clustering* tahun 2020. Pada Gambar 9 diketahui bahwa seluruh metode pada tahun 2020 memiliki nilai *silhouette score* tertinggi. Hal ini menunjukkan bahwa seluruh metode pada tahun 2020 merupakan hasil *clustering* yang optimum. Hasil *clustering* yang optimum akan dilakukan analisis karakteristiknya pada tahapan *deployment*. Analisis karakteristik dilakukan dengan menggunakan nilai *centroid*. Nilai *centroid* akan dikategorikan untuk melihat posisi dari masing – masing klaster. Penentuan rentang nilai kategori serta kategori yang digunakan terdapat pada bagian metodologi. Perhitungan rentang nilai dilakukan untuk masing – masing atribut.

Berdasarkan hasil *clustering* tahun 2020 yang terdapat pada Tabel 6, Tabel 7, Tabel 8, dan Tabel 9 diketahui bahwa anggota dari setiap klaster untuk hasil *clustering k-means* dan

hierarchical clustering adalah sama walaupun terjadi perbedaan pada letak klaster. Oleh karena itu hasil *clustering k-means* akan digunakan untuk dilakukan analisis karakteristiknya. Hal ini juga sesuai dengan nilai *centroid* yang sama untuk seluruh metode *clustering*. Berikut adalah nilai *centroid* seluruh metode, nilai *centroid k-means*, serta kategori masing – masing atribut yang dapat dilihat pada Tabel 11, Gambar 10, dan Tabel 12,.

Tabel 11 Nilai Centroid Metode Optimum

Metode	Centroid	Revenue Aero	Revenue Non - Aero	Pergerakan Penumpang	Pergerakan Pesawat
<i>K-Means</i>	Klaster 0	0,683	0,495	0,650	0,394
	Klaster 1	0,116	0,102	0,135	0,110
	Klaster 2	1,000	1,000	1,000	1,000
<i>Hierarchical Clustering "Single "</i>	Klaster 0	0,683	0,495	0,650	0,394
	Klaster 1	1,000	1,000	1,000	1,000
	Klaster 2	0,116	0,102	0,135	0,110
<i>Hierarchical Clustering "Complete "</i>	Klaster 0	0,116	0,102	0,135	0,110
	Klaster 1	0,683	0,495	0,650	0,394
	Klaster 2	1,000	1,000	1,000	1,000
<i>Hierarchical Clustering "Average "</i>	Klaster 0	0,683	0,495	0,650	0,394
	Klaster 1	1,000	1,000	1,000	1,000
	Klaster 2	0,116	0,102	0,135	0,110



Gambar 10 Radar Chart Nilai Centroid

Tabel 12 Kategori Atribut dan Klaster

	Revenue Aero	Revenue Non - Aero	Pergerakan Penumpang	Pergerakan Pesawat
Klaster 0	Sedang	Sedang	Sedang	Sedang
Klaster 1	Rendah	Rendah	Rendah	Rendah
Klaster 2	Tinggi	Tinggi	Tinggi	Tinggi

Berdasarkan pada Tabel 11 diketahui bahwa nilai *centroid* pada setiap klasternya adalah sama untuk seluruh atribut. Berdasarkan Gambar 10 diketahui bahwa klaster 2 memiliki nilai *centroid* terbesar dibandingkan klaster lainnya. Berdasarkan Tabel 12 diketahui bahwa seluruh atribut untuk klaster 0 berada pada kategori sedang, klaster 1 berada pada kategori rendah, dan klaster 2 berada pada kategori tinggi.

4. KESIMPULAN

Kesimpulan yang didapat pada penelitian ini adalah penerapan *clustering* dengan menggunakan metode *partitional (k-means)* dan *hierarchical clustering* menunjukkan bahwa hasil yang optimum didapatkan dari penerapan algoritma *k-means* dan *hierarchical clustering* untuk tahun 2020 dengan nilai *silhouette score* 0,678 dan memiliki tiga klaster tanpa adanya perbedaan anggota untuk setiap klasternya. Karakteristik dari hasil *clustering* yang optimum menunjukkan atribut *revenue aero*, *revenue non – aero*, pergerakan penumpang, dan

pergerakan pesawat berada pada kategori sedang untuk klaster 0, kategori rendah untuk klaster 1, dan kategori tinggi untuk klaster 2. Hasil *clustering* bandar udara dapat digunakan sebagai acuan dalam menentukan strategi pengelolaan bandar udara.

DAFTAR PUSTAKA

- Alasadi, S. A., & Bhaya, W. S. (2017). Review of Data Preprocessing Techniques in Data Mining. *Journal of Engineering and Applied Sciences*, *12*(16), 4102–4107.
- Alodhaibi, S., Burdett, R. L., & Yarlagadda, P. K. D. V. (2017). Framework for Airport Outbound Passenger Flow Modelling. *Procedia Engineering*, *174*, 1100–1109. <https://doi.org/10.1016/j.proeng.2017.01.263>
- H, R. K., Safwandi, S., Hasdyna, N., & Azizah, N. (2020). Analisis K-Means Clustering pada Data Sepeda Motor. *INFORMAL: Informatics Journal*, *5*(1), 10. <https://doi.org/10.19184/isj.v5i1.17071>
- Hanief, Y. N., & Purnomo, A. M. I. (2019). Menuju Pekan Olahraga Provinsi (PORPROV) Jawa Timur tahun 2019: Analisis kondisi fisik cabang olahraga petanque. *Prosiding Seminar Nasional IPTEK Olahraga (SENALOG)*, *2*(1), 1–4. <https://ejournal.unibabwi.ac.id/index.php/semnassenalog/article/view/575>
- Hazel, R. (2012). Guide to Airport performance measures. *Services Marketing Quarterly*, *33*(February), 55.
- Kansal, T., Bahuguna, S., Singh, V., & Choudhury, T. (2018). Customer Segmentation using K-means Clustering. *Proceedings of the International Conference on Computational Techniques, Electronics and Mechanical Systems, CTEMS 2018*, 135–139. <https://doi.org/10.1109/CTEMS.2018.8769171>
- Mangortey, E., Puranik, T. G., Pinon, O. J., & Mavris, D. N. (2020). Classification, analysis, and prediction of the daily operations of airports using machine learning. *AIAA Scitech 2020 Forum, 1 Part F*(January), 1–18. <https://doi.org/10.2514/6.2020-1196>
- Nasution, D. A., Khotimah, H. H., & Chamidah, N. (2019). Perbandingan Normalisasi Data untuk Klasifikasi Wine Menggunakan Algoritma K-NN. *Computer Engineering, Science and*

System Journal, 4(1), 78. <https://doi.org/10.24114/cess.v4i1.11458>

Orisa, M., & Faisal, A. (2021). Analisis Algoritma Partitioning Around Medoid untuk Penentuan Klusterisasi. *Jurnal Teknologi Informasi Dan Terapan (J-TIT)*, 8(2), 86–90.

Peraturan Menteri Perhubungan Nomor 69 Tahun 2013 Tentang Tata N Kebandarudaraan Nasional

Puspa, A. W. (2021). *Transformasi dan Konsistensi Bikin AP II Bisa Kelola 20 Bandara*. <https://ekonomi.bisnis.com/read/20210815/98/1429887/transformasi-dan-konsistensi-bikin-ap-ii-bisa-kelola-20-bandara>

Schröer, C., Kruse, F., & Gómez, J. M. (2021). A systematic literature review on applying CRISP-DM process model. *Procedia Computer Science*, 181(2019), 526–534. <https://doi.org/10.1016/j.procs.2021.01.199>

Setiani, B. (2015). *Prinsip-Prinsip Manajemen Pengelolaan Bandar Udara*. 3.

Thara, T. D. K., Prema, P. S., & Xiong, F. (2019). Auto-detection of epileptic seizure events using deep neural network with different feature scaling techniques. *Pattern Recognition Letters*, 128, 544–550. <https://doi.org/10.1016/j.patrec.2019.10.029> Undang – Undang Republik Indonesia Nomor 1 Tahun 2009 Tentang Penerbangan