BAB 4
HASIL PENELITIAN DAN PEMBAHASAN

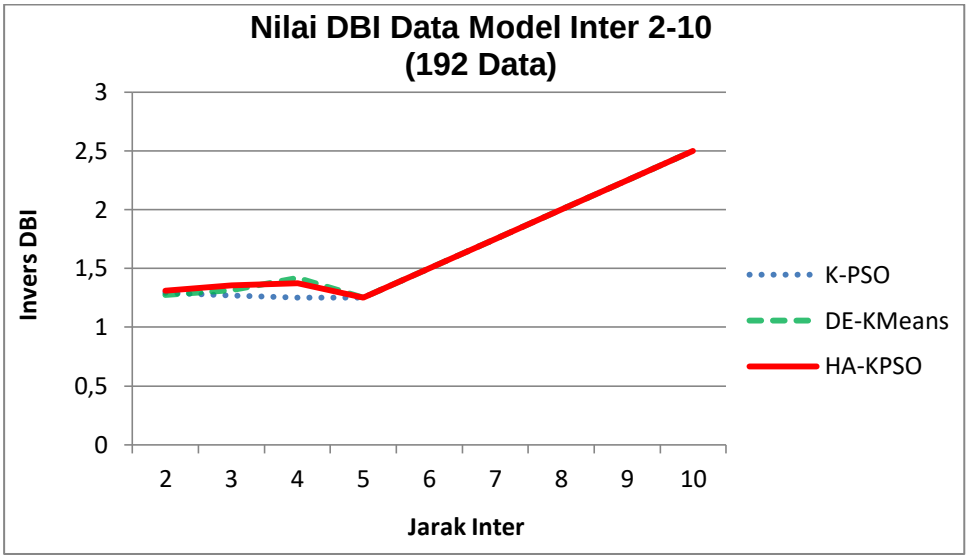
4.1 Hasil Pengujian Data Model

Pengujian data model diawali dengan mencari perilaku algoritma dalam membentuk hasil *clustering* untuk *dataset* buatan dengan jarak inter yang dekat dan yang jauh. Tabel 4.1 menunjukkan salah satu hasil percobaan yang dilakukan pada *dataset* 2 atribut – 2 *cluster* yang berjumlah 192 data untuk jarak inter 2 sampai 10.

Tabel 4.1 Nilai DBI *Dataset* buatan inter 2-10 (192 data)

Jarak Inter	Nilai DBI		
	K-PSO	DE-KMeans	HA-KPSO
2	0,77344	0,784793	0,7625
3	0,78645	0,7612	0,73623
4	0,79826	0,704198	0,72678
5	0,799832	0,799832	0,799832
6	0,666527	0,666527	0,666527
7	0,571309	0,571309	0,571309
8	0,499895	0,499895	0,499895
9	0,444351	0,444351	0,444351
10	0,399916	0,399916	0,399916

Tabel 4.1 memperlihatkan hasil validasi *cluster* DBI algoritma K-PSO, DE-KMeans, dan HA-KPSO. Untuk mempermudah perbandingan nilai DBI, hasil tersebut divisualisasikan ke dalam grafik yang dapat dilihat di Gambar 4.1



Gambar 4.1 Nilai DBI Data Model inter 2-10 (192 Data)

Pada Gambar 4.1 dapat dilihat bahwa pada *dataset* buatan dengan jarak inter di bawah 5, nilai DBI untuk algoritma K-PSO, DE-KMeans, dan HA-KPSO berbeda-beda, sementara pada jarak inter 5 hingga 10, ketiga algoritma memiliki nilai DBI yang sama. Hal ini menunjukkan bahwa pada saat jarak antar titik *centroid* berdekatan (bentuk *cluster* masih acak) ketiga algoritma menghasilkan kinerja yang berbeda-beda. Sementara saat jarak antar titik *centroid*-nya berjauhan, ketiga algoritma memiliki kinerja yang sama sebab bentuk *cluster* sudah terbentuk dengan baik sehingga jumlah *cluster* yang terdeteksi adalah benar yaitu 2 *cluster*. Berdasarkan hal tersebut maka untuk selanjutnya, pengujian difokuskan pada *dataset* buatan dengan jarak inter 2.

Pengujian data model ini dilakukan dengan menggunakan validasi *cluster* DBI terhadap hasil *clustering* algoritma K-PSO, DE-KMeans, dan HA-KPSO pada *dataset* buatan 2 dan 3 cluster dengan jumlah atribut dari 2 sampai 13 atribut dan *dataset* UCI yang terdiri dari *dataset* iris, wine dan diabetes. Hasil pengujian yang ditampilkan di bab 4 merupakan hasil yang dapat menunjukkan atau mewakili karakteristik algoritma yang diusulkan. Hasil pengujian terhadap data model yang tidak ditampilkan di bab ini dapat dilihat di lampiran A.

4.1.1 Pengujian *Dataset* Buatan 2 *Cluster*

Hasil Pengujian pada *dataset* buatan 2 *cluster* terdiri dari pengujian terhadap *dataset* 11 atribut, 12 atribut dan 13 atribut. Hasil pengujian atribut 2 s/d atribut 10 dapat dilihat di lampiran A.

1. Pengujian Pada *Dataset* 11 Atribut

Pengujian dilakukan sebanyak 10 kali menggunakan algoritma K-PSO, HA-KPSO dan DE-KMeans untuk dataset 11 atribut – 2 *cluster* dengan variasi jumlah data yang dimulai dari 24 data s/d 768 data. Hasil pengujian yang akan ditampilkan berupa rata-rata nilai validasi *cluster*, standar deviasi rata-rata validasi *cluster*, rata-wata waktu eksesusi dan standar deviasi rata-rata waktu eksekusi.

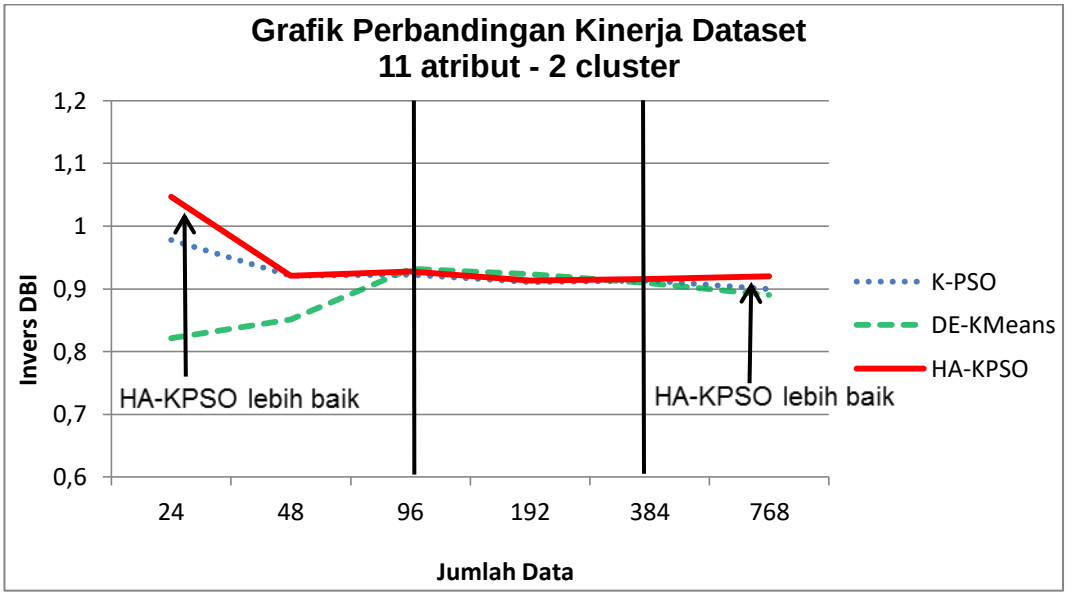
A. Hasil Pengujian Algoritma Bedasarkan nilai DBI

Pengujian hasil *clustering* mengunakan metode validasi *cluster* DBI pada algoritma K-PSO, HA-KPSO dan DE-KMeans untuk dataset 11 atribut – 2 *cluster* dapat dilihat Tabel 4.2.

Tabel 4.2 Hasil DBI dan Standar Deviasi *Dataset* 11 atribut – 2 *cluster*

Jumlah data	DBI			Standar Deviasi		
	K-PSO	DE-KMeans	HA-KPSO	K-PSO	DE-KMeans	HA-KPSO
24	1,02261957	1,217848	0,955166	0	0	0
48	1,08637064	1,174981	1,086370	0	0	0
96	1,08410463	1,073203	1,078385	0,00728	0	0,008658
192	1,09857005	1,082786	1,095548	0,01219	0	0,008147
384	1,09498928	1,100411	1,092215	0,03103	0	0,023609
768	1,1111376	1,1232461	1,086876	0,03323	0	0,025240

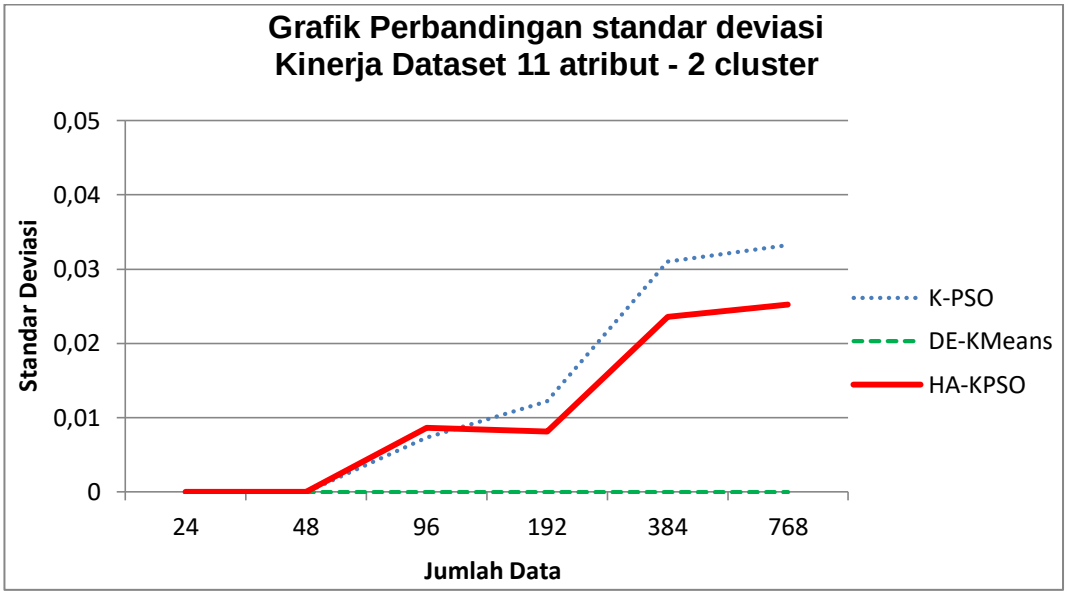
Tabel 4.2 menunjukkan rata-rata hasil pengujian DBI dan standar deviasi terhadap algoritma K-PSO, HA-KPSO dan DE-KMeans. Semakin kecil nilai DBI artinya semakin bagus kinerja algoritma. Sebaliknya semakin besar nilai DBI maka semakin buruk kinerja algoritma tersebut. Untuk memudahkan melihat perbandingan nilai DBI antar algoritma, maka hasil pengujian divisualisasikan ke dalam sebuah grafik yang ditampilkan di Gambar 4.2.



Gambar 4.2 Grafik Perbandingan DBI *Dataset* 11 atribut – 2 *cluster*

Berdasarkan grafik yang ditampilkan di Gambar 4.2 sumbu x merepresentasikan jumlah data dan sumbu y merepresentasikan nilai DBI. Untuk memudahkan membaca grafik maka nilai DBI pada Tabel 4.2 di*invers* terlebih dahulu sehingga semakin tinggi garis pada sumbu y artinya semakin baik kinerja algoritma. Sebaliknya semakin rendah garis pada sumbu y artinya semakin buruk kinerja algoritma. Dapat dilihat pada grafik bahwa Algoritma HA-KPSO memiliki kinerja yang lebih baik dibandingkan algoritma DE-KMeans dan K-PSO pada jumlah data < 100 data dan jumlah data > 400 data.

Selanjutnya, untuk melihat kepresisian algoritma maka dilakukan perbandingan terhadap standar deviasi dari rata-rata DBI yang dapat dilihat di Gambar 4.3.



Gambar 4.3 Grafik Perbandingan Standar Deviasi Kinerja *Dataset* 11 atribut – 2 *cluster*

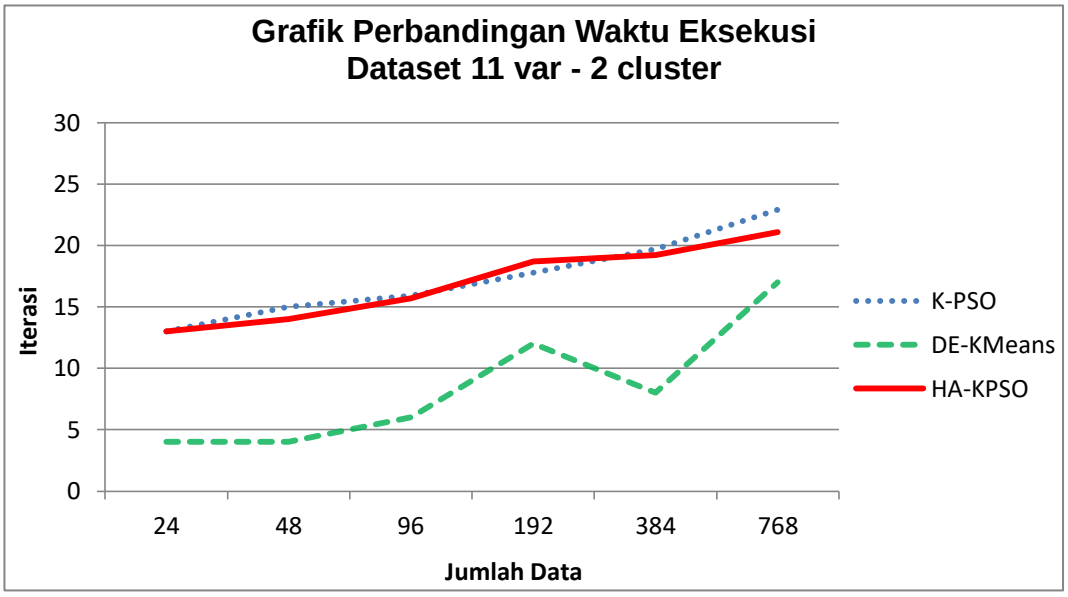
Berdasarkan grafik yang ditampilkan di Gambar 4.3 dapat dilihat bahwa nilai standar deviasi yang dimiliki algoritma HA-KPSO sama baiknya dengan K-PSO dan DE-KMeans pada jumlah data ≤ 48. Pada jumlah data > 48 data, nilai standar deviasi algoritma HA-KPSO menjadi lebih buruk dibandingkan DE-KMeans dan K-PSO. Namun standar deviasi dari HA-KPSO kembali menjadi lebih baik dibandingkan K-PSO pada jumlah data > 100 data. Hal ini menunjukkan bahwa algoritma HA-KPO mempunyai presisi yang lebih buruk dibandingkan dengan algoritma DE-KMeans dan lebih baik dibandingkan K-PSO saat jumlah data > 100 data.

Selanjutnya untuk mengetahui *cost* yang dibutuhkan dalam menghasilkan hasil *clustering*, dilakukan pengujian terhadap waktu eksekusi algoritma dengan melihat jumlah perulangan (*iterasi*) dari proses clustering algoritma tersebut. Waktu eksekusi algoritma K-PSO, HA-KPSO dan DE-KMeans pada *dataset* buatan 11 atribut – 2 *cluster* disajikan di Tabel 4.3.

Tabel 4.3 Waktu Eksekusi dan Standar Deviasi *Dataset* 11 atribut – 2 *cluster*

Jumlah data	Waktu Eksekusi			Standar Deviasi		
	K-PSO	DE-KMeans	HA-KPSO	K-PSO	DE-KMeans	HA-KPSO
24	13	4	13	13	4	13
48	15	4	14	15	4	14
96	15,9	6	15,7	15,9	6	15,7
192	17,8	12	18,7	17,8	12	18,7
384	19,7	8	19,2	19,7	8	19,2
768	22,9	17	21,1	22,9	17	21,1

Tabel 4.3 menunjukkan waktu eksekusi serta standar deviasi untuk dari algoritma K-PSO, HA-KPSO dan DE-KMeans. Semakin besar nilai iterasi artinya semakin lama waktu eksekusi algoritma dalam melakukan proses *clustering*. Sebaliknya semakin kecil nilai iterasi artinya semakin cepat waktu eksekusi algoritma dalam melakukan proses *clustering*. Untuk memudahkan melihat perbandingan waktu eksekusi dari ketiga algoritma tersebut maka hasil pengujian akan divisualisasikan ke dalam grafik yang ditampilkan di Gambar 4.4.

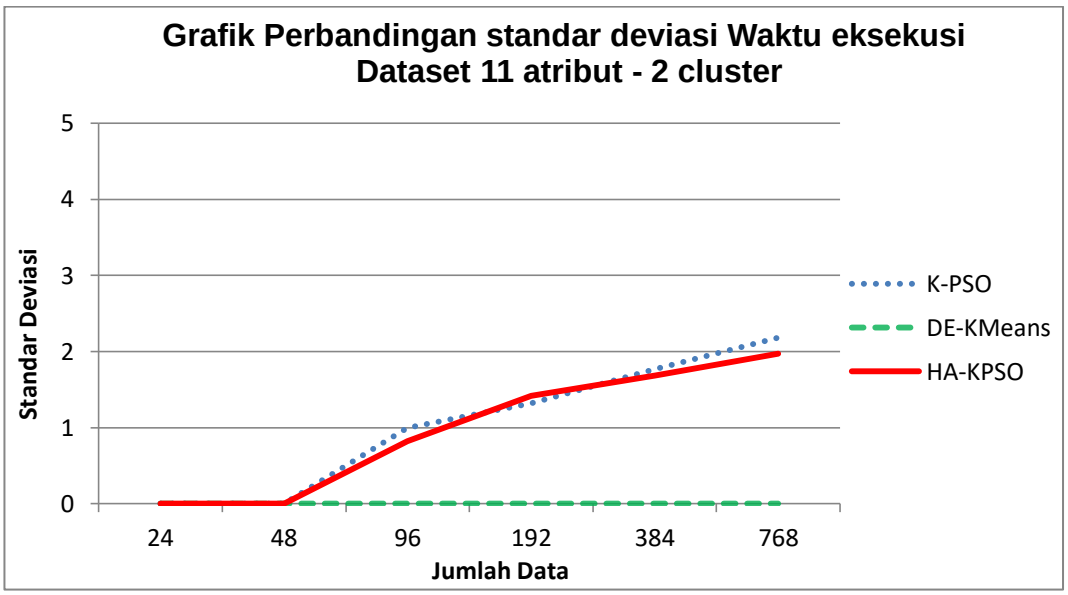


Gambar 4.4 Grafik Perbandingan Waktu Eksekusi *Dataset* 11 atribut – 2 *cluster*

Berdasarkan grafik yang ditampilkan di Gambar 4.4 sumbu x merepresentasikan jumlah data dan sumbu y merepresentasikan iterasi. Semakin

tinggi garis pada sumbu y artinya semakin lama waktu eksekusi algoritma. Sebaliknya semakin rendah garis pada sumbu y artinya semakin cepat waktu eksekusi algoritma. Dapat dilihat pada grafik bahwa algoritma HA-KPSO memiliki jumlah iterasi lebih sedikit dibandingkan algoritma K-PSO pada jumlah data < 100 data dan jumlah data > 300 data. Namun pada kondisi jumlah data yang sama, jumlah iterasi HA-KPSO lebih banyak dibandingkan DE-KMeans.

Selanjutnya untuk melihat kepresisian algoritma maka dilakukan perbandingan terhadap standar deviasi iterasinya yang dapat dilihat di Gambar 4.5.



Gambar 4.5 Grafik Perbandingan Standar Deviasi Waktu Eksekusi *Dataset* 11 atribut – 2 *cluster*

Berdasarkan grafik yang ditampilkan di Gambar 4.5 dapat dilihat pada *dataset* 11 atribut – 2 *cluster* algoritma HA-KPSO memiliki nilai standar deviasi lebih baik dibandingkan K-PSO pada jumlah data < 100 dan > dari 300 data. Namun pada kondisi ini jumlah iterasi dari algoritma HA-KPSO lebih buruk dibandingkan algoritma DE-KMeans. Hal ini menunjukkan bahwa algoritma HA-KPO mempunyai presisi yang lebih buruk dibandingkan dengan DE-KMeans dan lebih baik dari K-PSO pada jumlah data < 100 dan > 300 data.

2. Pengujian Data Model 12 Atribut

Pengujian dilakukan sebanyak 10 kali menggunakan algoritma K-PSO, HA-KPSO dan DE-KMeans untuk *dataset* 12 atribut – 2 *cluster* dengan variasi jumlah data yang dimulai dari 24 data s/d 768 data. Hasil pengujian yang akan ditampilkan berupa rata-rata nilai validasi *cluster*, standar deviasi rata-rata validasi *cluster*, rata-wata waktu eksesusi dan standar deviasi rata-rata waktu eksekusi.

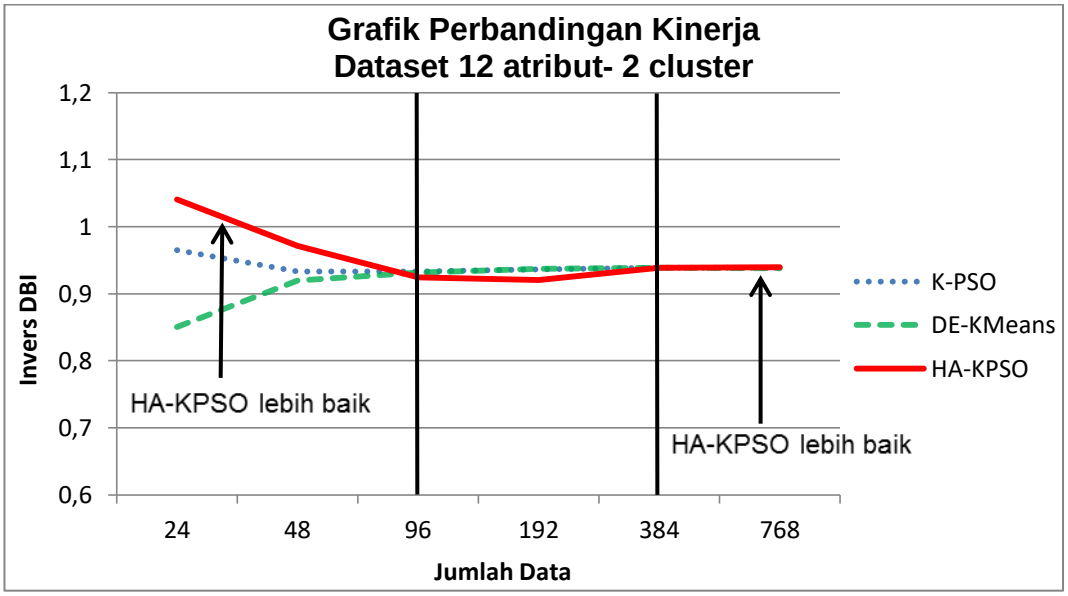
A. Hasil Pengujian Algoritma Bedasarkan nilai DBI

Pengujian hasil *clustering* mengunakan fungsi validasi *cluster* DBI pada algoritma K-PSO, HA-KPSO dan DE-KMeans pada dataset 12 atribut – 2 *cluster* dapat dilihat di Tabel 4.4.

Tabel 4.4 Hasil DBI dan Standar Deviasi *Dataset* 12 atribut – 2 *cluster*

Jumlah data	DBI			Standar Deviasi		
	K-PSO	DE-KMeans	HA-KPSO	K-PSO	DE-KMeans	HA-KPSO
24	1,035916738	1,175841	0,961051305	0	0	0
48	1,071987073	1,087088	1,029587668	0	0	0
96	1,071779078	1,073096	1,081868541	0,00266907	0	0,012423418
192	1,068143126	1,066708	1,08612534	0,001512337	0	0,015718015
384	1,065191765	1,0652897	1,065191765	0,000840042	0	0,015022792
768	1,065682199	1,0662272	1,064682199	0,00049054	0	0,017719232

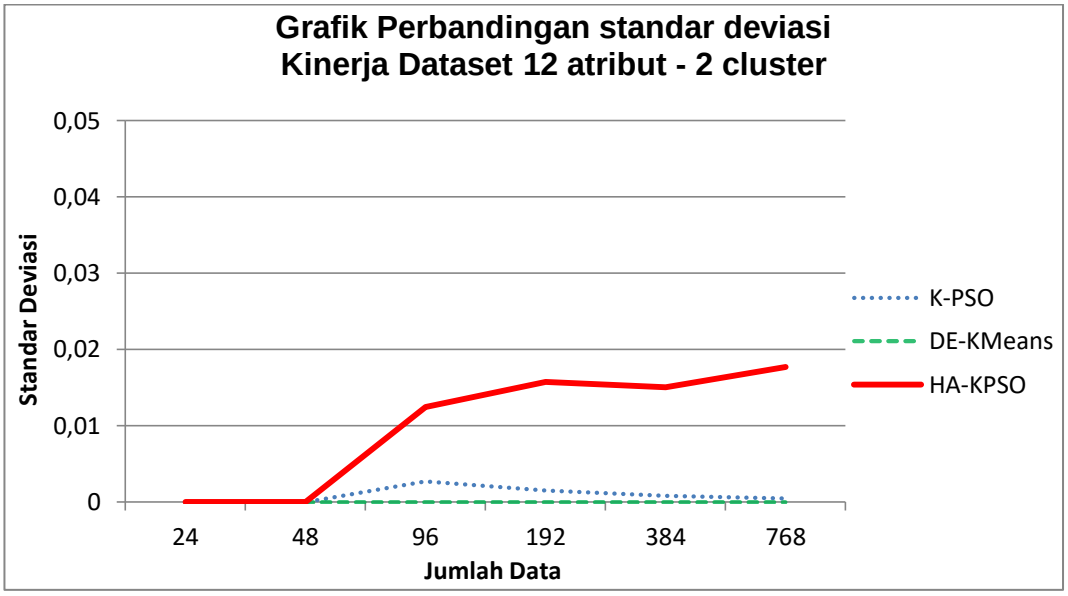
Tabel 4.4 menunjukkan rata-rata hasil pengujian DBI dan standar deviasi terhadap algoritma K-PSO, HA-KPSO dan DE-KMeans. Semakin kecil nilai DBI artinya semakin bagus kinerja suatu algoritma. Sebaliknya semakin besar nilai DBI maka semakin buruk kinerja algoritma tersebut. Untuk memudahkan melihat perbandingan nilai DBI antar algoritma, maka hasil pengujian divisualisasikan ke dalam sebuah grafik yang ditampilkan di Gambar 4.6.



Gambar 4.6 Grafik kinerja DBI data model 12 atribut – 2 *cluster*

Berdasarkan grafik yang ditampilkan di Gambar 4.6 sumbu x merepresentasikan jumlah data dan sumbu y merepresentasikan nilai DBI. Untuk memudahkan membaca grafik maka nilai DBI pada Tabel 4.4 di*invers* terlebih dahulu sehingga semakin tinggi garis pada sumbu y artinya semakin baik kinerja algoritma. Sebaliknya semakin rendah garis pada sumbu y artinya semakin buruk kinerja algoritma. Dapat dilihat pada grafik bahwa algoritma HA-KPSO memiliki kinerja yang lebih baik dibandingkan DE-KMeans dan K-PSO pada jumlah data < 100 data dan jumlah data > 400 data.

Selanjutnya, untuk melihat kepresisian algoritma maka dilakukan perbandingan terhadap standar deviasi dari rata-rata DBI yang dapat dilihat di Gambar 4.7.



Gambar 4.7 Grafik Standar Deviasi Kinerja *Dataset* 12 atribut – 2 *cluster*

Berdasarkan grafik yang ditampilkan di Gambar 4.7 dapat dilihat bahwa nilai standar deviasi dari algoritma HA-KPSO sama baiknya dibandingkan algoritma K-PSO dan DE-KMeans saat jumlah < 48 data. Sementara saat jumlah data > 48 data, Standar deviasi algoritma HA-KPSO menjadi lebih buruk dibandingkan K-PSO dan DE-KMeans. Hal ini menunjukkan bahwa algoritma HA-KPO mempunyai presisi yang lebih buruk dibandingkan dengan algoritma K-PSO DE-KMeans saat jumlah Data bertambah banyak.

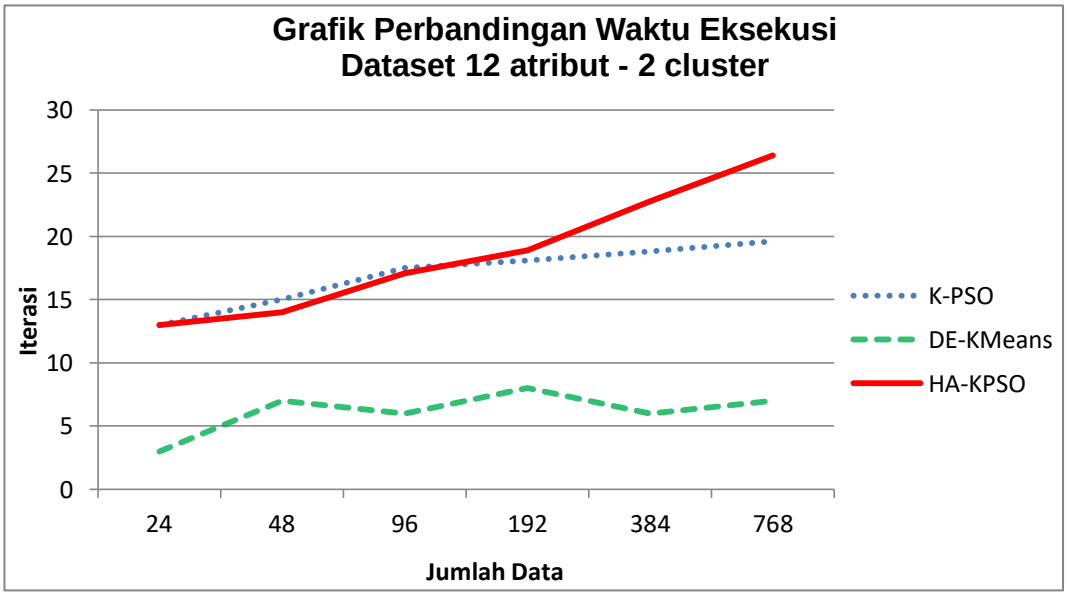
Selanjutnya untuk mengetahui *cost* yang dibutuhkan dalam menghasilkan hasil *clustering*, dilakukan pengujian terhadap waktu eksekusi algoritma dengan melihat jumlah perulangan (*iterasi*) dari proses clustering algoritma tersebut. Waktu eksekusi algoritma K-PSO, HA-KPSO dan DE-KMeans pada *dataset* 12 atribut – 2 *cluster* disajikan di Tabel 4.5

Tabel 4.5 Waktu Eksekusi dan Standar Deviasi *Dataset* 12 atribut – 2 cluster

Jumlah data	Waktu Eksekusi			Standar Deviasi		
	K-PSO	DE-KMeans	HA-KPSO	K-PSO	DE-KMeans	HA-KPSO
24	13	3	13	0	0	0
48	15	7	14	0	0	0
96	17,5	6	17,1	1,08012345	0	1,100504935
192	18,1	8	18,9	1,286683938	0	1,595131482

Jumlah data	Waktu Eksekusi			Standar Deviasi		
	K-PSO	DE-KMeans	HA-KPSO	K-PSO	DE-KMeans	HA-KPSO
384	18,8	6	22,8	1,619327707	0	1,988857852
768	19,6	7	26,4	1,95505044	0	2,065591118

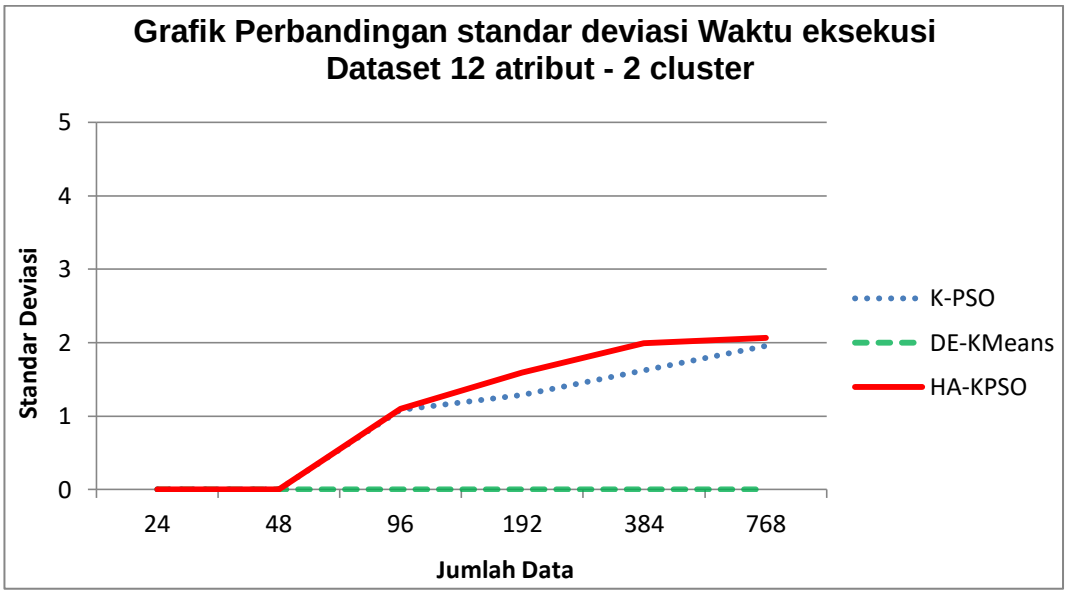
Tabel 4.5 menunjukkan waktu eksekusi serta standar deviasi untuk dari algoritma K-PSO, HA-KPSO dan DE-KMeans. Semakin besar nilai iterasi artinya semakin lama waktu eksekusi algoritma dalam melakukan proses *clustering*. Sebaliknya semakin kecil nilai iterasi artinya semakin cepat waktu eksekusi algoritma dalam melakukan proses *clustering*. Untuk memudahkan melihat perbandingan waktu eksekusi dari ketiga algoritma tersebut maka hasil pengujian akan divisualisasikan ke dalam grafik yang ditampilkan di Gambar 4.8.



Gambar 4.8 Grafik Waktu Eksekusi *Dataset* 12 atribut – 2 *cluster*

Berdasarkan grafik yang ditampilkan di Gambar 4.8, sumbu x merepresentasikan jumlah data dan sumbu y merepresentasikan iterasi. Semakin tinggi garis pada sumbu y artinya semakin lama waktu eksekusi algoritma. Sebaliknya semakin rendah garis pada sumbu y artinya semakin cepat waktu eksekusi algoritma. Dapat dilihat algoritma HA-KPSO memiliki jumlah iterasi lebih sedikit dibandingkan algoritma K-PSO pada jumlah data <100 data. Namun

pada kondisi ini jumlah iterasi HA-KPSO lebih banyak dibandingkan algoritma DE-KMeans. Sedangkan pada jumlah data ≥ 100 data, jumlah iterasi algoritma HA-KPSO lebih banyak dibandingkan algoritma K-PSO dan DE-KMeans. Selanjutnya, untuk melihat kepresisian algoritma maka dilakukan perbandingan standar deviasi iterasinya yang dapat dilihat di Gambar 4.9.



Gambar 4.9 Grafik Standar Deviasi Waktu Eksekusi *Dataset* 12 atribut – 2 *cluster*

Berdasarkan grafik yang ditampilkan di Gambar 4.9 dapat dilihat bahwa nilai standar deviasi algoritma HA-KPSO sama baiknya jika dibandingkan dengan algoritma K-PSO dan DE-KMeans pada jumlah data ≤ 48 data. Sementara pada jumlah data > 100 data, standar deviasi algoritma HA-KPSO menjadi lebih buruk dibandingkan algorima K-PSO dan DE-KMeans.

3. Pengujian *Dataset* 13 Atribut

Pengujian dilakukan sebanyak 10 kali menggunakan algoritma K-PSO, HA-KPSO dan DE-KMeans untuk *dataset* 13 atribut – 2 *cluster* dengan variasi jumlah data yang dimulai dari 24 data s/d 768 data. Hasil pengujian yang akan ditampilkan berupa rata-rata nilai validasi *cluster*, standar deviasi rata-rata

validasi *cluster*, rata-wata waktu eksesusi dan standar deviasi rata-rata waktu eksekusi.

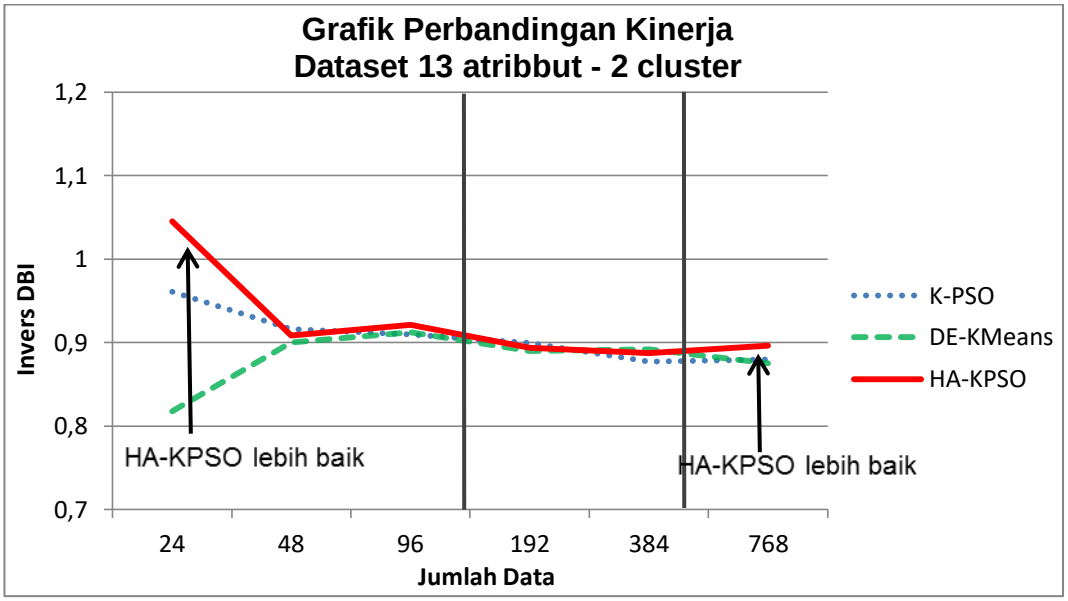
A. Hasil Pengujian Algoritma Bedasarkan nilai DBI

Pengujian hasil *clustering* mengunakan metode validasi *cluster* DBI pada algoritma K-PSO, HA-KPSO dan DE-KMeans untuk dataset 13 atribut – 2 *cluster* dilihat di Tabel 4.6.

Tabel 4.6 Hasil DBI dan Standar Deviasi *Dataset* 13 atribut – 2 *cluster*

Jumlah data	DBI			Standar Deviasi		
	K-PSO	DE-KMeans	HA-KPSO	K-PSO	DE-KMeans	HA-KPSO
24	1,040817828	1,222762	0,956358429	0	0	0
48	1,091575444	1,110717	1,100674027	0	0	0
96	1,09494715	1,095814	1,085548049	0,004774165	0	0,006028845
192	1,111434293	1,123383	1,118826627	0,034231722	0	0,028122162
384	1,13965848	1,1209246	1,127050047	0,031925382	0	0,025503494
768	1,136824992	1,1421315	1,115351814	0,043598702	0	0,038617848

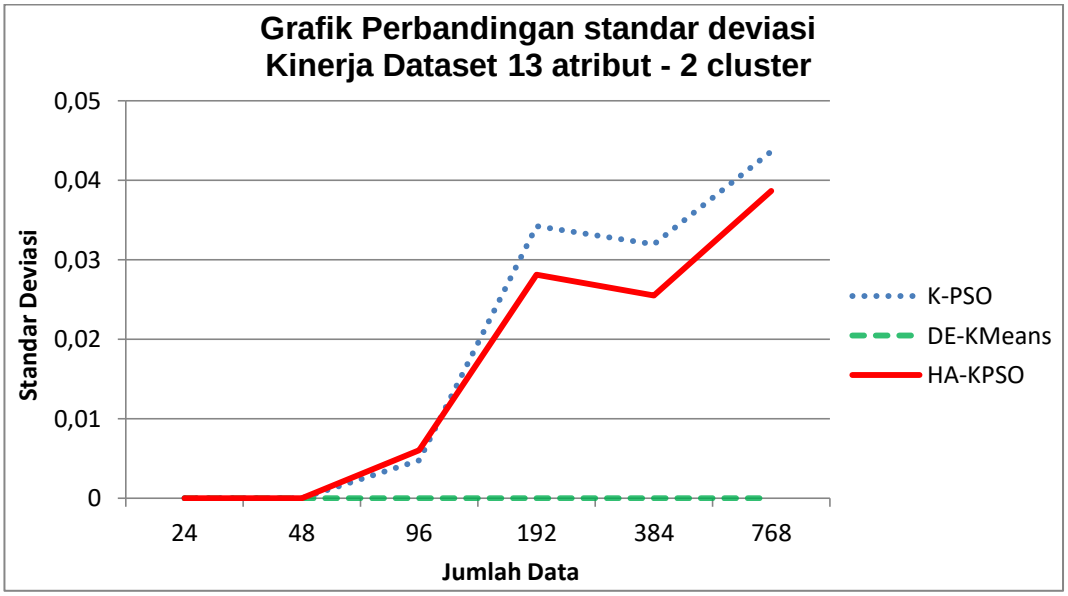
Tabel 4.6 menunjukkan rata-rata hasil pengujian DBI dan standar deviasi terhadap algoritma K-PSO, HA-KPSO dan DE-KMeans. Semakin kecil nilai DBI artinya semakin bagus kinerja algoritma. Sebaliknya semakin besar nilai DBI maka semakin buruk kinerja algoritma tersebut. Untuk memudahkan melihat perbandingan nilai DBI antar algoritma, maka hasil pengujian divisualisasikan ke dalam sebuah grafik yang ditampilkan di Gambar 4.10.



Gambar 4.10 Grafik kinerja DBI *Dataset* 13 atribut – 2 *cluster*

Berdasarkan grafik yang ditampilkan di Gambar 4.10 sumbu x merepresentasikan jumlah data dan sumbu y merepresentasikan nilai DBI. Untuk memudahkan membaca grafik maka nilai DBI pada Tabel 4.6 di*invers* terlebih dahulu sehingga semakin tinggi garis pada sumbu y artinya semakin baik kinerja algoritma. Sebaliknya semakin rendah garis pada sumbu y artinya semakin buruk kinerja algoritma. Dapat dilihat pada grafik bahwa algoritma HA-KPSO memiliki kinerja yang lebih baik dibandingkan algoritma DE-KMeans dan K-PSO pada jumlah data < 200 dan jumlah data > 500 data.

Selanjutnya, untuk melihat kepresisian algoritma maka dilakukan perbandingan terhadap standar deviasi dari rata-rata DBI yang dapat dilihat di Gambar 4.11.



Gambar 4.11 Grafik Standar Deviasi Kinerja *Dataset* 13 atribut – 2 *cluster*

Berdasarkan grafik yang ditampilkan di Gambar 4.11 dapat dilihat nilai standar deviasi algoritma HA-KPSO sama baiknya dengan K-PSO dan DE-KMeans pada jumlah data ≤ 48. Namun pada jumlah data > 48 data, Nilai standar deviasi algoritma HA-KPSO menjadi lebih buruk dibandingkan algoritma DE-KMeans dan K-PSO. Pada jumlah data >100 data, nilai standar deviasi dari algoritma HA-KPSO kembali menjadi lebih baik dibandingkan algoritma K-PSO namun tetap lebih buruk dibandingkan algoritma DE-KMeans.

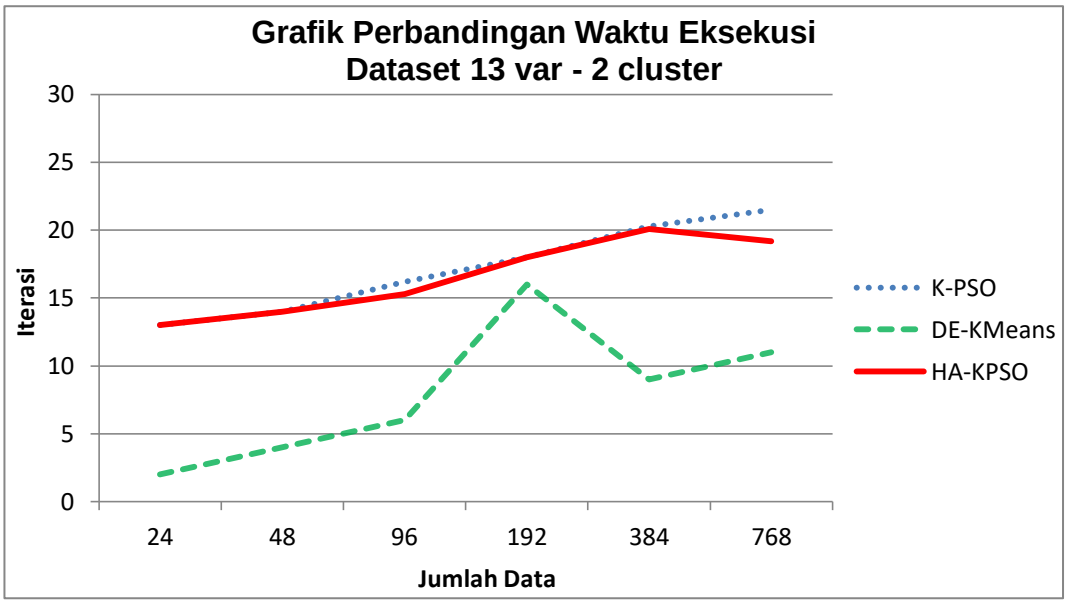
Selanjutnya untuk mengetahui *cost* yang dibutuhkan dalam menghasilkan hasil *clustering*, dilakukan pengujian terhadap waktu eksekusi algoritma dengan melihat jumlah perulangan (*iterasi*) dari proses clustering algoritma tersebut. Waktu eksekusi algoritma K-PSO, HA-KPSO dan DE-KMeans *dataset* 12 atribut – 2 *cluster* disajikan di Tabel 4.7

Tabel 4.7 Waktu Eksekusi dan Standar Deviasi *Dataset* 13 atribut – 2 cluster

Jumlah data	Waktu Eksekusi			Standar Deviasi		
	K-PSO	DE-KMeans	HA-KPSO	K-PSO	DE-KMeans	HA-KPSO
24	13	2	13	0	0	0
48	14	4	14	0	0	0
96	16,2	6	15,3	1,032795559	0	0,948683298
192	18	16	18	1,699673171	0	1,490711985

Jumlah data	Waktu Eksekusi			Standar Deviasi		
	K-PSO	DE-KMeans	HA-KPSO	K-PSO	DE-KMeans	HA-KPSO
384	20,3	9	20,1	1,946506843	0	1,728840331
768	21,5	11	19,2	2,121320344	0	1,932183566

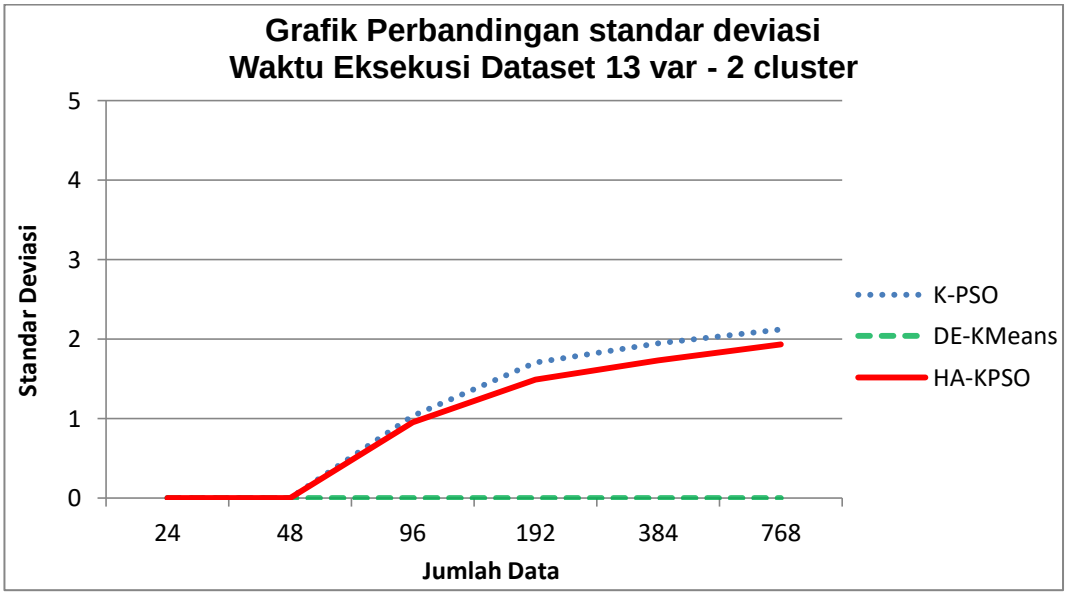
Tabel 4.7 menunjukkan waktu eksekusi serta standar deviasi untuk dari algoritma K-PSO, HA-KPSO dan DE-KMeans. Semakin besar nilai iterasi artinya semakin lama waktu eksekusi algoritma dalam melakukan proses *clustering*. Sebaliknya semakin kecil nilai iterasi artinya semakin cepat waktu eksekusi algoritma dalam melakukan proses *clustering*. Untuk memudahkan melihat perbandingan waktu eksekusi dari ketiga algoritma tersebut maka hasil pengujian akan divisualisasikan ke dalam grafik yang ditampilkan di Gambar 4.12.



Gambar 4.12 Grafik Waktu Eksekusi *Dataset* 13 atribut – 2 *cluster*

Berdasarkan grafik yang ditampilkan di Gambar 4.12 sumbu x merepresentasikan jumlah data dan sumbu y merepresentasikan iterasi. Semakin tinggi garis pada sumbu y artinya semakin lama waktu eksekusi algoritma. Sebaliknya semakin rendah garis pada sumbu y artinya semakin cepat waktu eksekusi algoritma. Dapat dilihat algoritma HA-KPSO memiliki jumlah iterasi yang lebih sedikit dibandingkan algoritma K-PSO pada jumlah data < 48 data.

Jika dibandingan dengan algoritma DE-KMeans, Algoritma HA-KPSO selalu mempunyai jumlah iterasi yang lebih banyak di seluruh variasi jumlah data yang diuji. Selanjutnya, untuk melihat kepresisian algoritma maka dilakukan perbandingan standar deviasi iterasinya yang dapat dilihat di Gambar 4.13.



Gambar 4.13 Grafik Standar Deviasi Waktu Eksekusi *Dataset* 13 atribut – 2 *cluster*

Berdasarkan grafik yang ditampilkan di Gambar 4.13 dapat dilihat standar deviasi dari algoritma HA-KPSO sama baiknya dibandingkan algoritma K-PSO dan DE-KMeans pada jumlah data ≤ 48 data. Pada jumlah data > 100 data, Standar deviasi algoritma HA-KPSO menjadi lebih baik dibandingkan algoritma K-PSO. Hal ini menunjukkan bahwa algoritma HA-KPO mempunyai presisi lebih buruk dibandingkan dengan DE-KMeans dan presisi lebih baik dibandingkan K-PSO pada jumlah data > 100.

4.1.2 Pengujian *Dataset Buatan* 3 *Cluster*

Hasil Pengujian pada *dataset* buatan 3 *cluster* terdiri dari pengujian terhadap *dataset* 11 atribut, 12 atribut dan 13 atribut. Hasil pengujian untuk atribut 2 s/d atribut 10 dapat dilihat di lampiran A.

1. Pengujian *Dataset* 11 Atribut

Pengujian dilakukan sebanyak 10 kali menggunakan algoritma K-PSO, HA-KPSO dan DE-KMeans untuk *dataset* 11 atribut – 3 *cluster* dengan variasi jumlah data yang dimulai dari 36 data s/d 1152 data. Hasil pengujian yang akan ditampilkan berupa rata-rata nilai validasi *cluster*, standar deviasi rata-rata validasi *cluster*, rata-wata waktu eksesusi dan standar deviasi rata-rata waktu eksekusi.

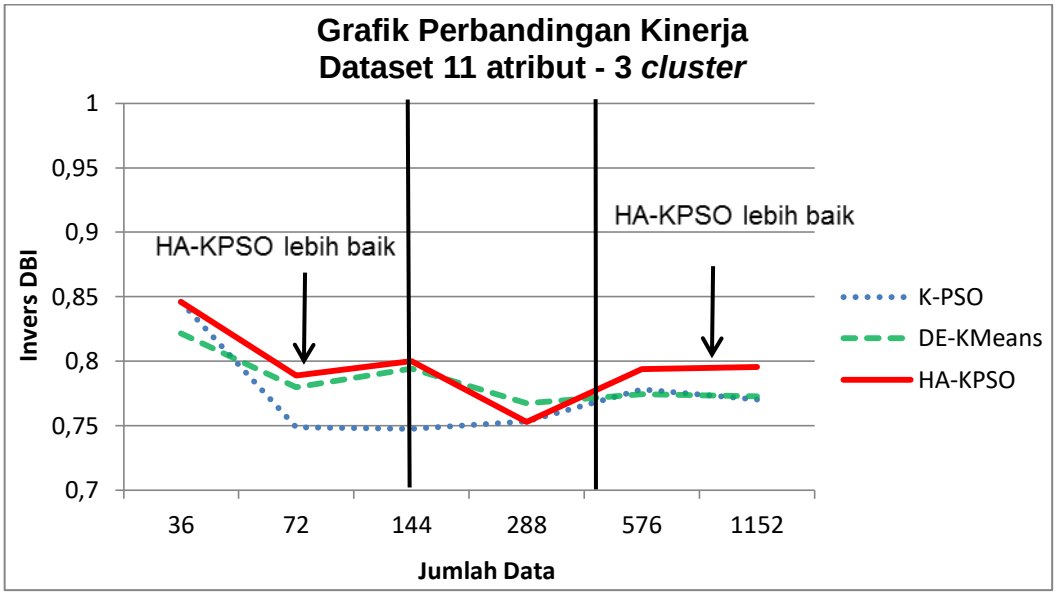
A. Hasil Pengujian Algoritma Bedasarkan nilai DBI

Pengujian hasil *clustering* mengunakan metode validasi *cluster* DBI pada algoritma K-PSO, HA-KPSO dan DE-KMeans untuk dataset 11 atribut – 3 *cluster* dilihat di Tabel 4.8.

Tabel 4.8 Hasil DBI dan Standar Deviasi *Dataset* 11 atribut – 3 *cluster*

Jumlah data	DBI			Standar Deviasi		
	K-PSO	DE-KMeans	HA-KPSO	K-PSO	DE-KMeans	HA-KPSO
36	1,1821741	1,217134917	1,182174196	0	0	0
72	1,3357004	1,282352054	1,267293158	0	0	0
144	1,3378594	1,308646967	1,249759165	0,01120569	0	0,0028135
228	1,3274931	1,252902391	1,327908583	0,01373370	0	0,0135279
576	1,2850002	1,291197254	1,259634323	0,02198583	0	0,0001319
1152	1,2979069	1,294241882	1,257204453	0,00330271	0	0,0004934

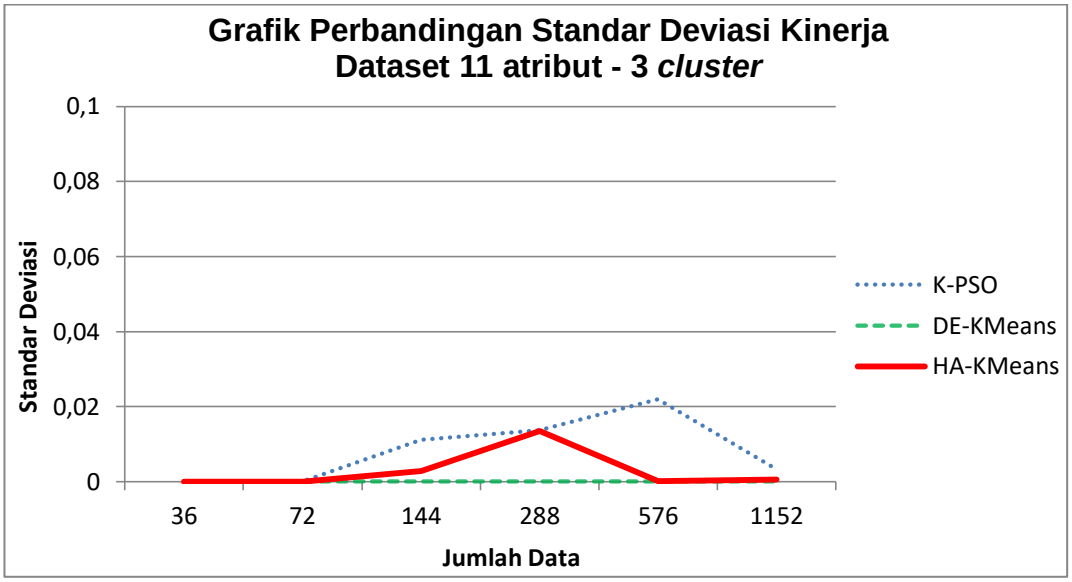
Tabel 4.8 menunjukkan rata-rata hasil pengujian DBI dan standar deviasi terhadap algoritma K-PSO, HA-KPSO dan DE-KMeans. Semakin kecil nilai DBI artinya semakin bagus kinerja algoritma. Sebaliknya semakin besar nilai DBI maka semakin buruk kinerja algoritma tersebut. Untuk memudahkan melihat perbandingan nilai DBI antar algoritma, maka hasil pengujian divisualisasikan ke dalam sebuah grafik yang ditampilkan di Gambar 4.14.



Gambar 4.14 Grafik kinerja DBI *Dataset* 11 atribut – 3 *cluster*

Berdasarkan grafik yang ditampilkan di Gambar 4.14 sumbu x merepresentasikan jumlah data dan sumbu y merepresentasikan nilai DBI. Untuk memudahkan membaca grafik maka nilai DBI pada Tabel 4.8 di*invers* terlebih dahulu sehingga semakin tinggi garis pada sumbu y artinya semakin baik kinerja algoritma. Sebaliknya semakin rendah garis pada sumbu y artinya semakin buruk kinerja algoritma. Dapat dilihat Algoritma HA-KPSO memiliki kinerja yang selalu lebih baik pada jumlah data < 144 data dan jumlah data > 400 dibandingkan dengan algoritma K-PSO dan DE-Kmeans.

Selanjutnya, untuk melihat kepresisian algoritma maka dilakukan perbandingan terhadap standar deviasi dari rata-rata DBI yang dapat dilihat di Gambar 4.15.



Gambar 4.15 Grafik Standar Deviasi Kinerja *Dataset* 11 atribut – 3 *cluster*

Berdasarkan grafik yang ditampilkan di Gambar 4.15 secara keseluruhan standar deviasi dari algoritma HA-KPSO lebih baik dibandingkan algoritma K-PSO di seluruh jumah data yang diuji. Sementara jika dibandingkan dengan algoritma DE-Kmeans, Algoritma HA-KPSO memiliki standar deviasi yang lebih buruk di seluruh jumlah data yang di uji. Hal ini menunjukkan algoritma HA-KPO mempunyai presisi yang lebih baik dari algoritma K-PSO namun mempunyai presisi yang lebih buruk dibandingkan DE-KMeans.

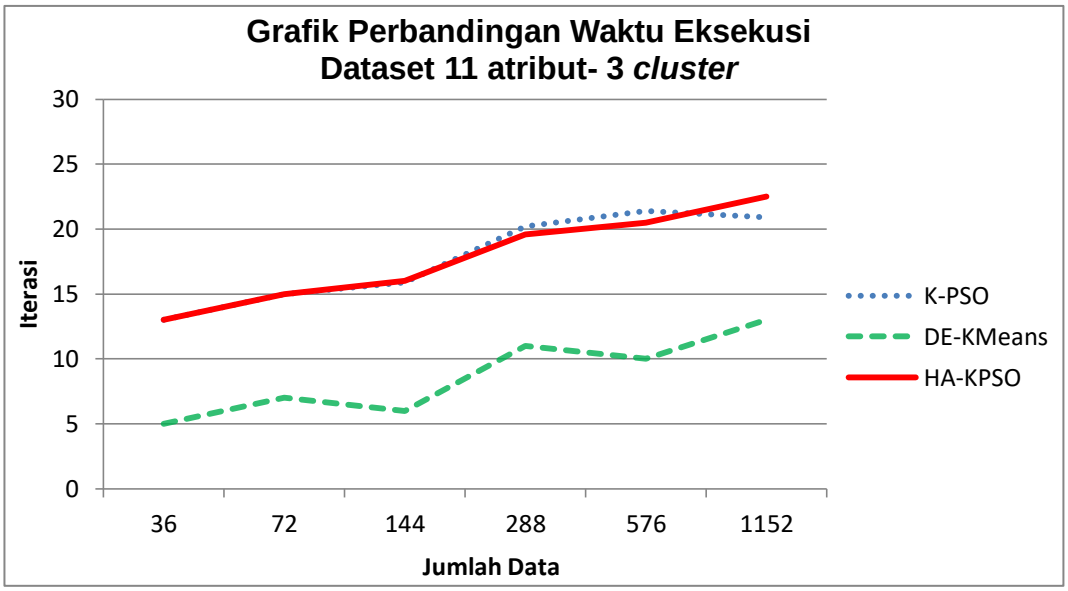
Selanjutnya untuk mengetahui *cost* yang dibutuhkan dalam menghasilkan hasil *clustering*, dilakukan pengujian terhadap waktu eksekusi algoritma dengan melihat jumlah perulangan (*iterasi*) dari proses clustering algoritma tersebut. Waktu eksekusi algoritma K-PSO, HA-KPSO dan DE-KMeans *dataset* 11 atribut - 3 *cluster* disajikan di Tabel 4.9.

Tabel 4.9 Waktu Eksekusi dan Standar Deviasi *Dataset* 11 atribut – 3 *cluster*

Jumlah data	Waktu Eksekusi			Standar Deviasi		
	K-PSO	DE-KMeans	HA-KPSO	K-PSO	DE-KMeans	HA-KPSO
36	13	5	13	0	0	0
72	15	7	15	0	0	0
144	15,9	6	16	1,100504935	0	0,699205899

Jumlah data	Waktu Eksekusi			Standar Deviasi		
	K-PSO	DE-KMeans	HA-KPSO	K-PSO	DE-KMeans	HA-KPSO
228	20,2	11	19,6	1,751190072	0	1,429840706
576	21,4	10	20,5	1,776388346	0	1,702938637
1152	20,9	13	22,5	2,233582076	0	1,900292375

Tabel 4.9 menunjukkan waktu eksekusi serta standar deviasi untuk dari algoritma K-PSO, HA-KPSO dan DE-KMeans. Semakin besar nilai iterasi artinya semakin lama waktu eksekusi algoritma dalam melakukan proses *clustering*. Sebaliknya semakin kecil nilai iterasi artinya semakin cepat waktu eksekusi algoritma dalam melakukan proses *clustering*. Untuk memudahkan melihat perbandingan waktu eksekusi dari ketiga algoritma tersebut maka hasil pengujian akan divisualisasikan ke dalam grafik yang ditampilkan di Gambar 4.16

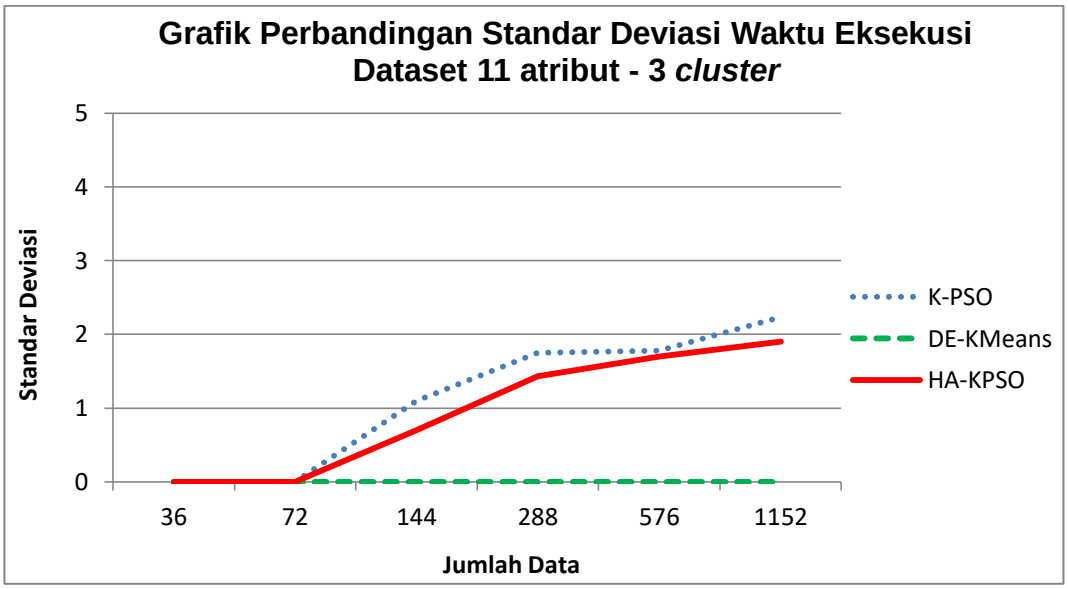


Gambar 4.16 Grafik Waktu Eksekusi *Dataset* 11 atribut – 3 *cluster*

Berdasarkan grafik yang ditampilkan di Gambar 4.16 sumbu x merepresentasikan jumlah data dan sumbu y merepresentasikan iterasi. Semakin tinggi garis pada sumbu y artinya semakin lama waktu eksekusi algoritma. Sebaliknya semakin rendah garis pada sumbu y artinya semakin cepat waktu eksekusi algoritma. Dapat algoritma HA-KPSO secara keseluruhan memiliki

jumlah iterasi yang lebih sedikit dibandingkan algoritma K-PSO di seluruh jumlah data yang diuji. Sementara jika dibandingkan dengan algoritma DE-KMeans, algoritma HA-KPSO mempunyai jumlah iterasi yang lebih banyak.

Selanjutnya, untuk melihat kepresisian algoritma maka dilakukan perbandingan standar deviasi iterasinya yang dapat dilihat di Gambar 4.17.



Gambar 4.17 Grafik Standar Deviasi Waktu Eksekusi *Dataset* 11 atribut – 3 *cluster*

Berdasarkan grafik yang ditampilkan di Gambar 4.17 algoritma HA-KPSO memiliki standar deviasi iterasi yang selalu lebih buruk dibandingkan algoritma DE-KMeans pada jumlah data ≥ 72. Sedangkan jika dibandingkan dengan algoritma K-PSO, standar deviasi algoritma HA-KPSO selalu lebih baik diseluruh jumlah data yang diuji.

2. Pengujian *Dataset* 12 Atribut

Pengujian dilakukan sebanyak 10 kali menggunakan algoritma K-PSO, HA-KPSO dan DE-KMeans untuk *dataset* 12 atribut – 3 *cluster* dengan variasi jumlah data yang dimulai dari 36 data s/d 1152 data. Hasil pengujian yang akan ditampilkan berupa rata-rata nilai validasi *cluster*, standar deviasi rata-rata

validasi *cluster*, rata-wata waktu eksesusi dan standar deviasi rata-rata waktu eksekusi.

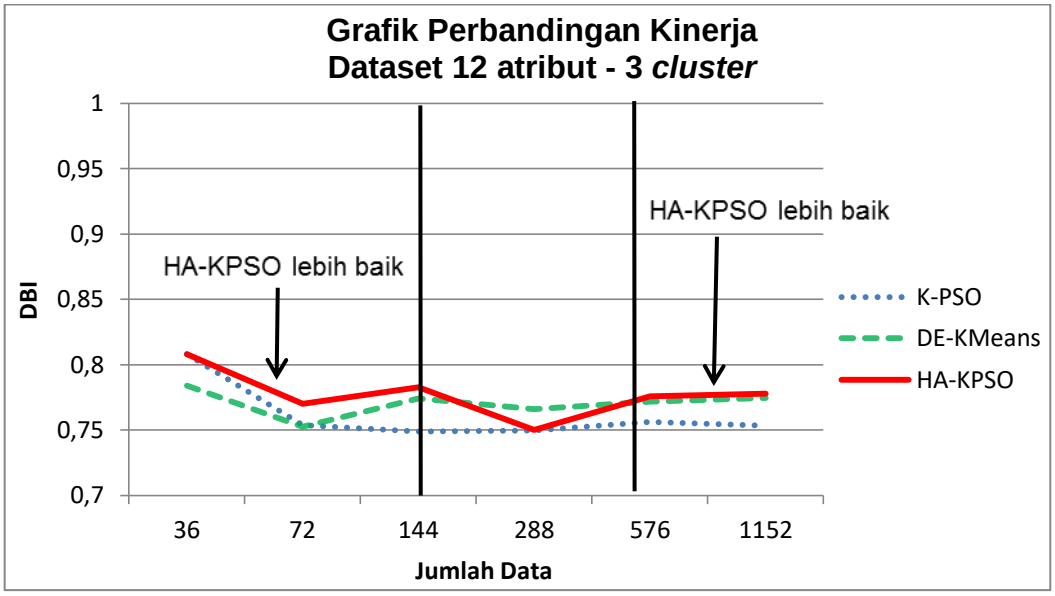
A. Hasil Pengujian Algoritma Bedasarkan nilai DBI

Pengujian hasil *clustering* mengunakan metode validasi *cluster* DBI pada algoritma K-PSO, HA-KPSO dan DE-KMeans untuk dataset 12 atribut – 3 *cluster* dilihat di Tabel 4.10.

Tabel 4.10 Hasil DBI dan Standar Deviasi *Dataset* 12 atribut – 3 *cluster*

Jumlah data	DBI			Standar Deviasi		
	K-PSO	DE-KMeans	HA-KPSO	K-PSO	DE-KMeans	HA-KPSO
36	1,2377787	1,27547398	1,237778774	0	0	0
72	1,3266157	1,329141786	1,298572308	0	0	0
144	1,3353103	1,291753356	1,27781415	0,00909824	0	0,0003343
228	1,3335815	1,305588184	1,333213385	0,01553872	0	0,0156858
576	1,3222048	1,295964686	1,289058724	0,00209912	0	0,0031618
1152	1,3271513	1,291151394	1,285721951	0,00135730	0	0,0009660

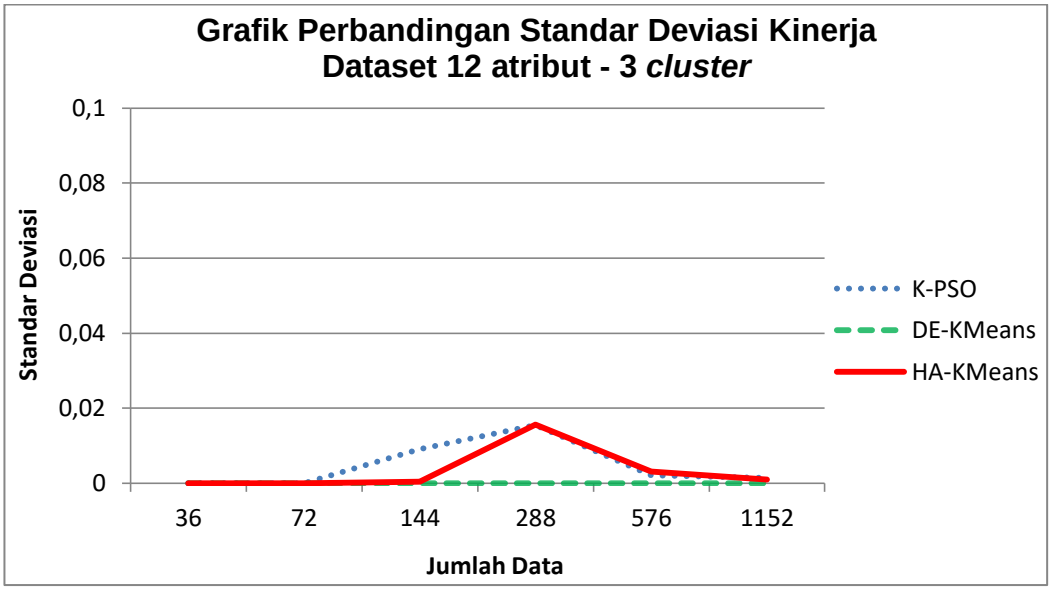
Tabel 4.10 menunjukkan rata-rata hasil pengujian DBI dan standar deviasi terhadap algoritma K-PSO, HA-KPSO dan DE-KMeans. Semakin kecil nilai DBI artinya semakin bagus kinerja algoritma. Sebaliknya semakin besar nilai DBI maka semakin buruk kinerja algoritma tersebut. Untuk memudahkan melihat perbandingan nilai DBI antar algoritma, maka hasil pengujian divisualisasikan ke dalam sebuah grafik yang ditampilkan di Gambar 4.18**.**



Gambar 4.18 Grafik kinerja DBI *Dataset* 12 atribut – 3 *cluster*

Berdasarkan grafik yang ditampilkan di Gambar 4.18 sumbu x merepresentasikan jumlah data dan sumbu y merepresentasikan nilai DBI. Untuk memudahkan membaca grafik maka nilai DBI pada Tabel 4.12 di*invers* terlebih dahulu sehingga semakin tinggi garis pada sumbu y artinya semakin baik kinerja algoritma. Sebaliknya semakin rendah garis pada sumbu y artinya semakin buruk kinerja algoritma. Dapat dilihat Algoritma HA-KPSO memiliki kinerja yang lebih baik pada jumlah data < 144 data dan jumlah data > 500 data dibandingkan algoritma K-PSO dan DE-Kmeans.

Selanjutnya, untuk melihat kepresisian algoritma maka dilakukan perbandingan terhadap standar deviasi dari rata-rata DBI yang dapat dilihat di Gambar 4.19.



Gambar 4.19 Grafik Standar Deviasi Kinerja *Dataset* 12 atribut – 3 *cluster*

Berdasarkan grafik yang ditampilkan di Gambar 4.19 secara keseluruhan standar deviasi dari algoritma HA-KPSO lebih baik dibandingkan algoritma K-PSO di seluruh jumah data yang diuji. Sementara jika dibandingkan dengan algoritma DE-Kmeans, Algoritma HA-KPSO memiliki standar deviasi yang lebih buruk di seluruh jumlah data yang di uji. Hal ini menunjukkan algoritma HA-KPO mempunyai presisi yang lebih baik dari algoritma K-PSO namun mempunyai presisi yang lebih buruk dibandingkan DE-KMeans.

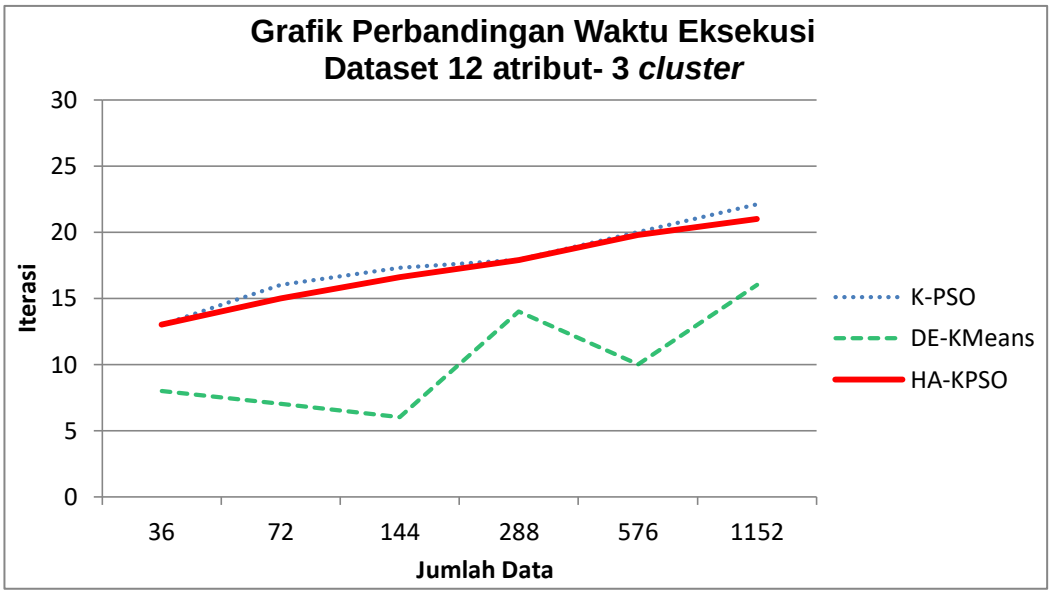
Selanjutnya untuk mengetahui *cost* yang dibutuhkan dalam menghasilkan hasil *clustering*, dilakukan pengujian terhadap waktu eksekusi algoritma dengan melihat jumlah perulangan (*iterasi*) dari proses clustering algoritma tersebut. Waktu eksekusi algoritma K-PSO, HA-KPSO dan DE-KMeans *dataset* 12 atribut - 3 *cluster* disajikan di Tabel 4.11

Tabel 4.11 Waktu Eksekusi dan Standar Deviasi *Dataset* 12 atribut – 3 *cluster*

Jumlah data	Waktu Eksekusi			Standar Deviasi		
	K-PSO	DE-KMeans	HA-KPSO	K-PSO	DE-KMeans	HA-KPSO
36	13	8	13	0	0	0
72	16	7	15	0	0	0
144	17,3	6	16,6	0,948683298	0	0,843274043

Jumlah data	Waktu Eksekusi			Standar Deviasi		
	K-PSO	DE-KMeans	HA-KPSO	K-PSO	DE-KMeans	HA-KPSO
228	17,9	14	17,9	1,370320319	0	1,370320319
576	20	10	19,8	1,56347192	0	1,549193338
1152	22,1	16	21	2,330951165	0	2

Tabel 4.11 menunjukkan waktu eksekusi serta standar deviasi untuk dari algoritma K-PSO, HA-KPSO dan DE-KMeans. Semakin besar nilai iterasi artinya semakin lama waktu eksekusi algoritma dalam melakukan proses *clustering*. Sebaliknya semakin kecil nilai iterasi artinya semakin cepat waktu eksekusi algoritma dalam melakukan proses *clustering*. Untuk memudahkan melihat perbandingan waktu eksekusi dari ketiga algoritma tersebut maka hasil pengujian akan divisualisasikan ke dalam grafik yang ditampilkan di Gambar 4.20.

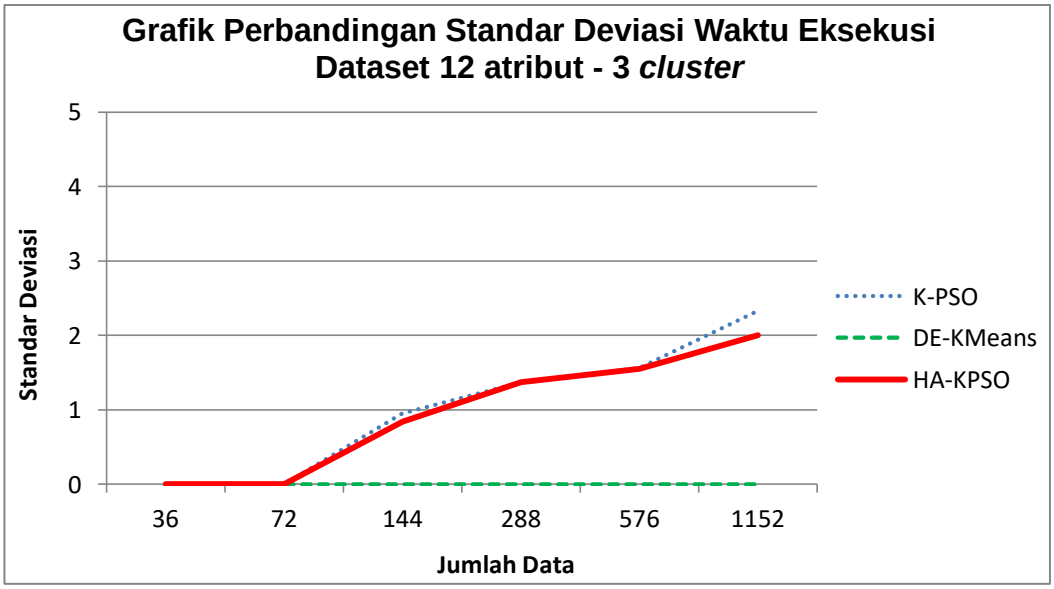


Gambar 4.20 Grafik Waktu Eksekusi *Dataset* 12 atribut – 3 *cluster*

Berdasarkan grafik yang ditampilkan di Gambar 4.21 sumbu x merepresentasikan jumlah data dan sumbu y merepresentasikan iterasi. Semakin tinggi garis pada sumbu y artinya semakin lama waktu eksekusi algoritma. Sebaliknya semakin rendah garis pada sumbu y artinya semakin cepat waktu eksekusi algoritma. Pada jumlah data < 500 Algoritma HA-KPSO memiliki

jumlah iterasi lebih sedikit dibandingkan algoritma K-PSO. Pada jumlah > 500 algoritma HA-KPSO memiliki jumlah iterasi lebih banyak dari K-PSO. Sementara jika dibandingkan dengan algoritma DE-KMeans, algoritma HA-KPSO mempunyai jumlah iterasi lebih banyak.

Selanjutnya, untuk melihat kepresisian algoritma maka dilakukan perbandingan standar deviasi iterasinya yang dapat dilihat di Gambar 4.21.



Gambar 4.21 Grafik Standar Deviasi Waktu Eksekusi *Dataset* 12 atribut – 3 *cluster*

Berdasarkan grafik yang ditampilkan di Gambar 4.25 algoritma HA-KPSO memiliki standar deviasi iterasi yang selalu lebih buruk dibandingkan algoritma DE-KMeans pada jumlah data ≥ 72. Sedangkan jika dibandingkan dengan algoritma K-PSO, standar deviasi algoritma HA-KPSO selalu lebih baik diseluruh jumlah data yang diuji.

3. Pengujian *Dataset* 13 Atribut

Pengujian dilakukan sebanyak 10 kali menggunakan algoritma K-PSO, HA-KPSO dan DE-KMeans untuk *dataset* 13 atribut – 3 *cluster* dengan variasi jumlah data yang dimulai dari 36 data s/d 1152 data. Hasil pengujian yang akan ditampilkan berupa rata-rata nilai validasi *cluster*, standar deviasi rata-rata

validasi *cluster*, rata-wata waktu eksesusi dan standar deviasi rata-rata waktu eksekusi.

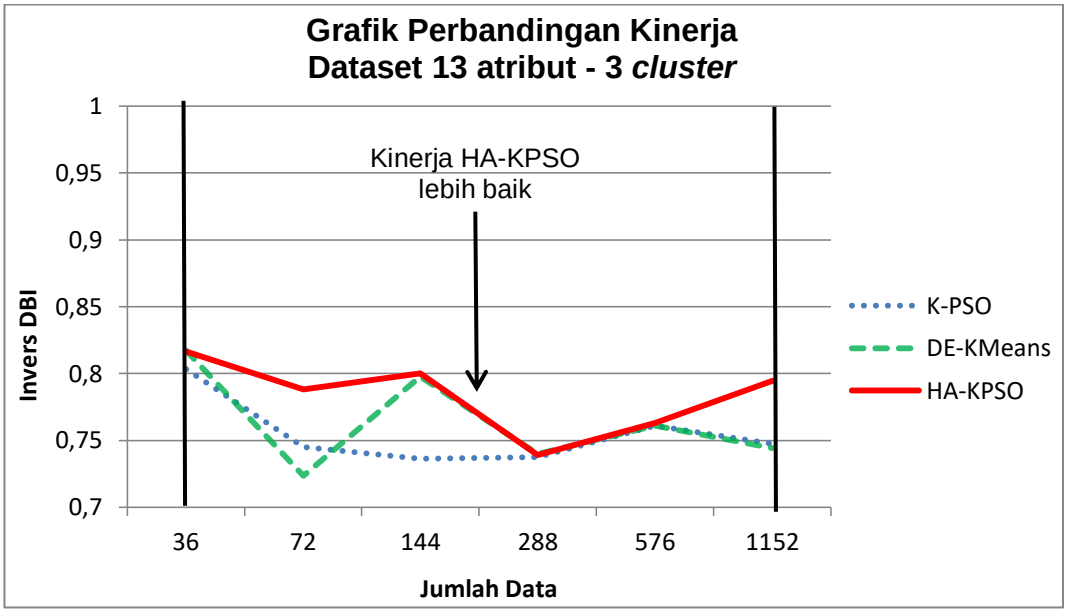
A. Hasil Pengujian Algoritma Bedasarkan nilai DBI

Pengujian hasil *clustering* mengunakan metode validasi *cluster* DBI pada algoritma K-PSO, HA-KPSO dan DE-KMeans untuk dataset 13 atribut – 3 *cluster* dilihat di Tabel 4.12.

Tabel 4.12 Hasil DBI dan Standar Deviasi *Dataset* 13 atribut – 3 *cluster*

Jumlah data	DBI			Standar Deviasi		
	K-PSO	DE-KMeans	HA-KPSO	K-PSO	DE-KMeans	HA-KPSO
36	1,2446555	1,2238752	1,224655584	0	0	0
72	1,3422511	1,3818867	1,269154279	0,02682048	0	0
144	1,3580203	1,2536848	1,249779701	0,01577842	0	0,0026947
228	1,3561074	1,3515024	1,353194264	0,01319749	0	0,0140482
576	1,3138185	1,3141423	1,31056883	0,00383285	0	0,0152104
1152	1,3376043	1,3436678	1,258587919	0,02536641	0	0

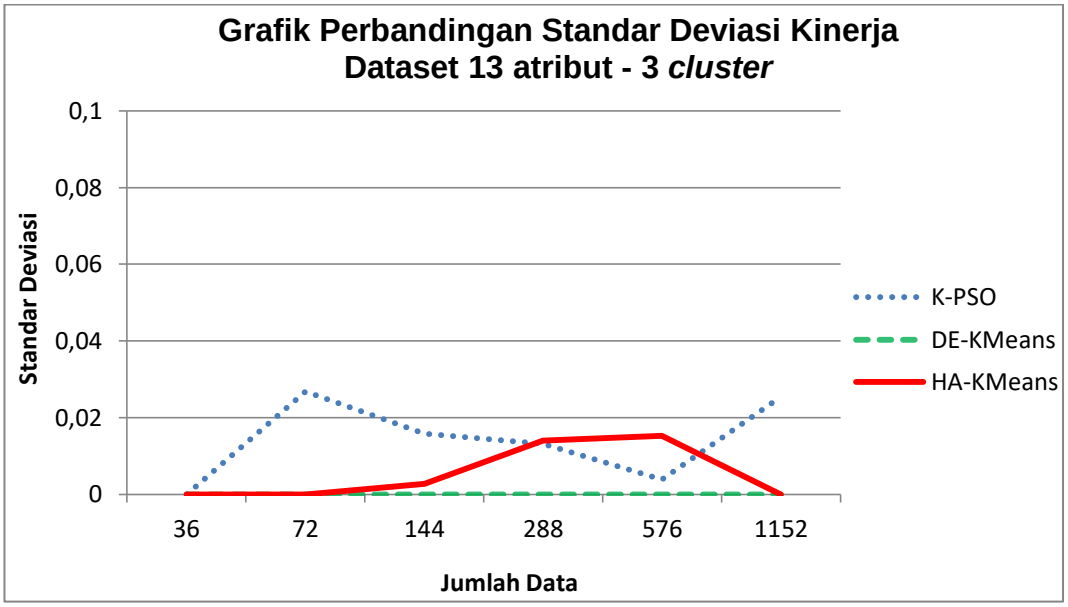
Tabel 4.12 menunjukkan rata-rata hasil pengujian DBI dan standar deviasi terhadap algoritma K-PSO, HA-KPSO dan DE-KMeans. Semakin kecil nilai DBI artinya semakin bagus kinerja algoritma. Sebaliknya semakin besar nilai DBI maka semakin buruk kinerja algoritma tersebut. Untuk memudahkan melihat perbandingan nilai DBI antar algoritma, maka hasil pengujian divisualisasikan ke dalam sebuah grafik yang ditampilkan di Gambar 4.22.



Gambar 4.22 Grafik kinerja DBI *Dataset* 13 atribut – 3 *cluster*

Berdasarkan grafik yang ditampilkan di Gambar 4.22 sumbu x merepresentasikan jumlah data dan sumbu y merepresentasikan nilai DBI. Untuk memudahkan membaca grafik maka nilai DBI pada Tabel 4.12 di*invers* terlebih dahulu sehingga semakin tinggi garis pada sumbu y artinya semakin baik kinerja algoritma. Sebaliknya semakin rendah garis pada sumbu y artinya semakin buruk kinerja algoritma. Dapat dilihat bahwa algoritma HA-KPSO memiliki kinerja yang lebih baik dibandingkan algoritma K-PSO dan DE-Kmeans diseluruh jumlah data yang diuji pada dataset 13 atribut.

Selanjutnya, untuk melihat kepresisian algoritma maka dilakukan perbandingan terhadap standar deviasi dari rata-rata DBI yang dapat dilihat di Gambar 4.23.



Gambar 4.23 Grafik Standar Deviasi Kinerja *Dataset* 13 atribut – 3 *cluster*

Berdasarkan grafik yang ditampilkan di Gambar 4.23 secara keseluruhan standar deviasi dari algoritma HA-KPSO lebih baik dibandingkan algoritma K-PSO di seluruh jumah data yang diuji. Sementara jika dibandingkan dengan algoritma DE-Kmeans, Algoritma HA-KPSO memiliki standar deviasi yang lebih buruk di seluruh jumlah data yang di uji. Hal ini menunjukkan algoritma HA-KPO mempunyai presisi yang lebih baik dari algoritma K-PSO namun mempunyai presisi yang lebih buruk dibandingkan DE-KMeans.

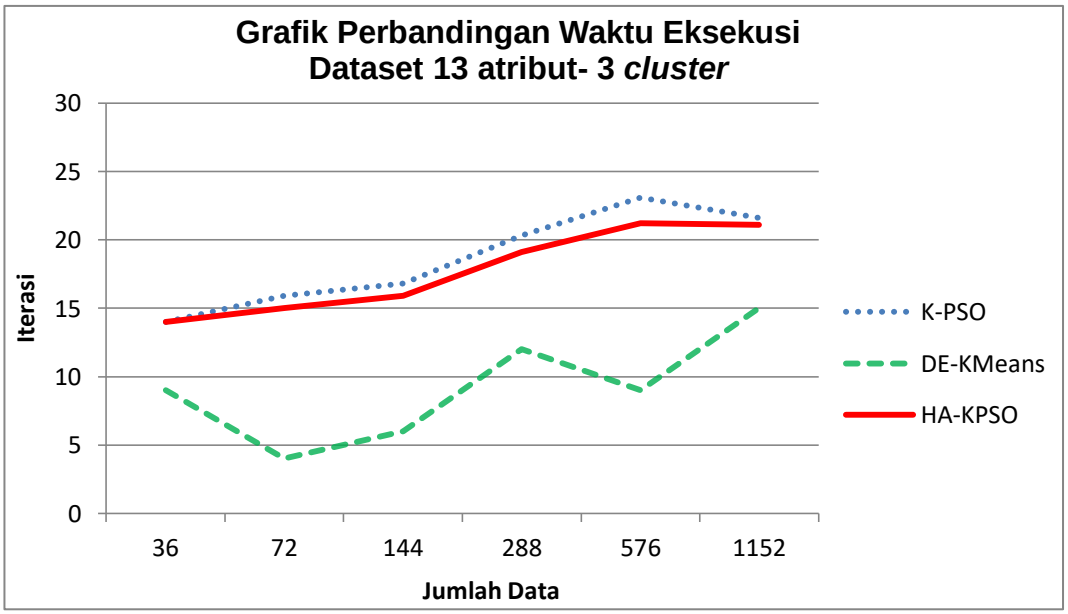
Selanjutnya untuk mengetahui *cost* yang dibutuhkan dalam menghasilkan hasil *clustering*, dilakukan pengujian terhadap waktu eksekusi algoritma dengan melihat jumlah perulangan (*iterasi*) dari proses clustering algoritma tersebut. Waktu eksekusi algoritma K-PSO, HA-KPSO dan DE-KMeans *dataset* 13 atribut – 3 *cluster* disajikan di Tabel 4.13.

Tabel 4.13 Waktu Eksekusi dan Standar Deviasi *Dataset* 13 atribut – 3 *cluster*

Jumlah data	Waktu Eksekusi			Standar Deviasi		
	K-PSO	DE-KMeans	HA-KPSO	K-PSO	DE-KMeans	HA-KPSO
36	14	9	14	0	0	0
72	15,9	4	15	0,875595036	0	0
144	16,8	6	15,9	1,751190072	0	1,449137675

Jumlah data	Waktu Eksekusi			Standar Deviasi		
	K-PSO	DE-KMeans	HA-KPSO	K-PSO	DE-KMeans	HA-KPSO
228	20,3	12	19,1	2,162817093	0	2,183269719
576	23,1	9	21,2	2,378141198	0	1,988857852
1152	21,6	15	21,1	2,67498702	0	1,969207398

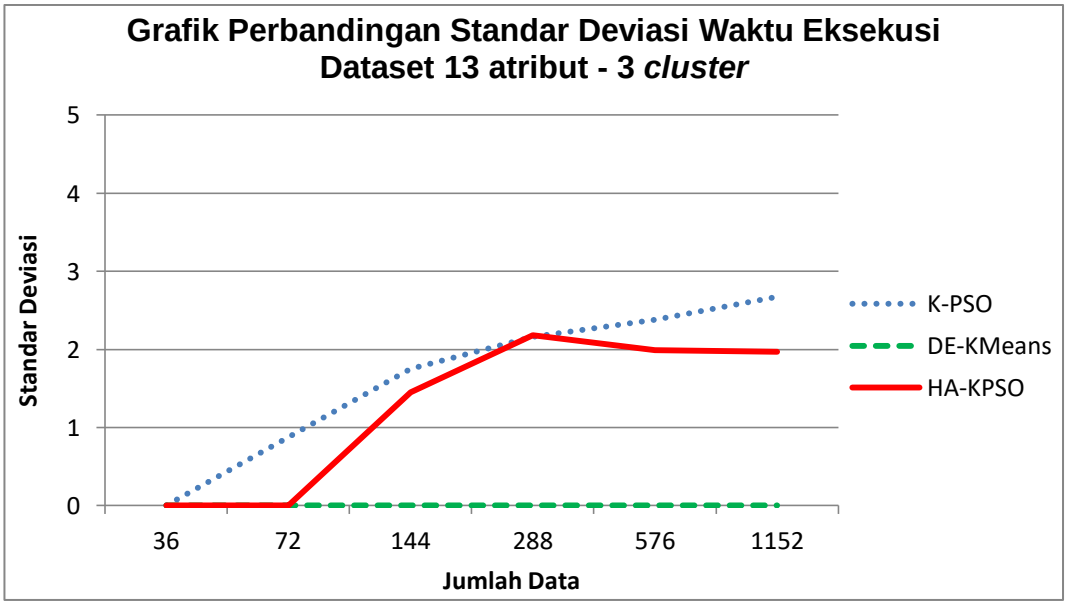
Tabel 4.13 menunjukkan waktu eksekusi serta standar deviasi untuk dari algoritma K-PSO, HA-KPSO dan DE-KMeans. Semakin besar nilai iterasi artinya semakin lama waktu eksekusi algoritma dalam melakukan proses *clustering*. Sebaliknya semakin kecil nilai iterasi artinya semakin cepat waktu eksekusi algoritma dalam melakukan proses *clustering*. Untuk memudahkan melihat perbandingan waktu eksekusi dari ketiga algoritma tersebut maka hasil pengujian akan divisualisasikan ke dalam grafik yang ditampilkan di Gambar 4.24.



Gambar 4.24 Grafik Waktu Eksekusi *Dataset* 13 atribut – 3 *cluster*

Berdasarkan grafik yang ditampilkan di Gambar 4.24 sumbu x merepresentasikan jumlah data dan sumbu y merepresentasikan iterasi. Semakin tinggi garis pada sumbu y artinya semakin lama waktu eksekusi algoritma. Sebaliknya semakin rendah garis pada sumbu y artinya semakin cepat waktu

eksekusi algoritma.. Dapat algoritma HA-KPSO secara keseluruhan memiliki jumlah iterasi yang lebih sedikit dibandingkan algoritma K-PSO di seluruh jumlah data yang diuji. Sementara jika dibandingkan dengan algoritma DE-KMeans, algoritma HA-KPSO mempunyai jumlah iterasi yang lebih banyak. Selanjutnya, untuk melihat kepresisian algoritma maka dilakukan perbandingan standar deviasi iterasinya yang dapat dilihat di Gambar 4.25.



Gambar 4.25 Grafik Standar Deviasi Waktu Eksekusi *Dataset* 13 atribut – 3 *cluster*

Berdasarkan grafik yang ditampilkan di Gambar 4.25 algoritma HA-KPSO memiliki standar deviasi iterasi yang selalu lebih buruk dibandingkan algoritma DE-KMeans pada jumlah data ≥ 72. Sedangkan jika dibandingkan dengan algoritma K-PSO, standar deviasi algoritma HA-KPSO selalu lebih baik diseluruh jumlah data yang diuji.

4.1.3 Pengujian *Dataset* Iris

Pengujian *dataset* iris dilakukan sebanyak 10 kali menggunakan algoritma K-PSO, HA-KPSO dan DE-KMeans dengan variasi jumlah data yang dimulai dari 150 data s/d 768 data. Hasil pengujian yang akan ditampilkan berupa rata-rata

nilai validasi *cluster*, standar deviasi rata-rata validasi *cluster*, rata-wata waktu eksesusi dan standar deviasi rata-rata waktu eksekusi.

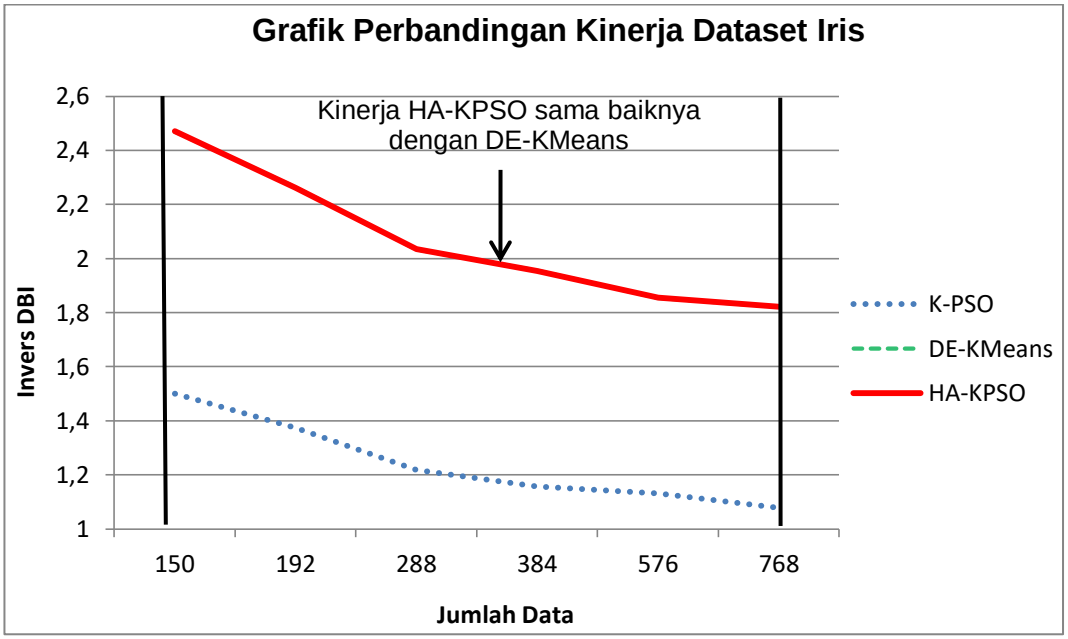
A. Hasil Pengujian Algoritma Bedasarkan nilai DBI

Pengujian hasil *clustering* mengunakan metode validasi *cluster* DBI pada algoritma K-PSO, HA-KPSO dan DE-KMeans untuk *dataset* iris dapat dilihat di Tabel 4.14.

Tabel 4.14 Hasil DBI dan Standar Deviasi Data Iris

Jumlah data	DBI			Standar Deviasi		
	K-PSO	DE-KMeans	HA-KPSO	K-PSO	DE-KMeans	HA-KPSO
150	0,666391211	0,4048341	0,404834136	0	0	0
192	0,727482014	0,442028	0,442028035	0,00086372	0	0
288	0,821139737	0,4915204	0,491520368	0,004112615	0	0
368	0,864962736	0,5117116	0,511711632	0,000617106	0	0
576	0,884315044	0,5389612	0,538958525	0,000450349	0	0
768	0,927491663	0,5490954	0,549095395	0,004440444	0	0

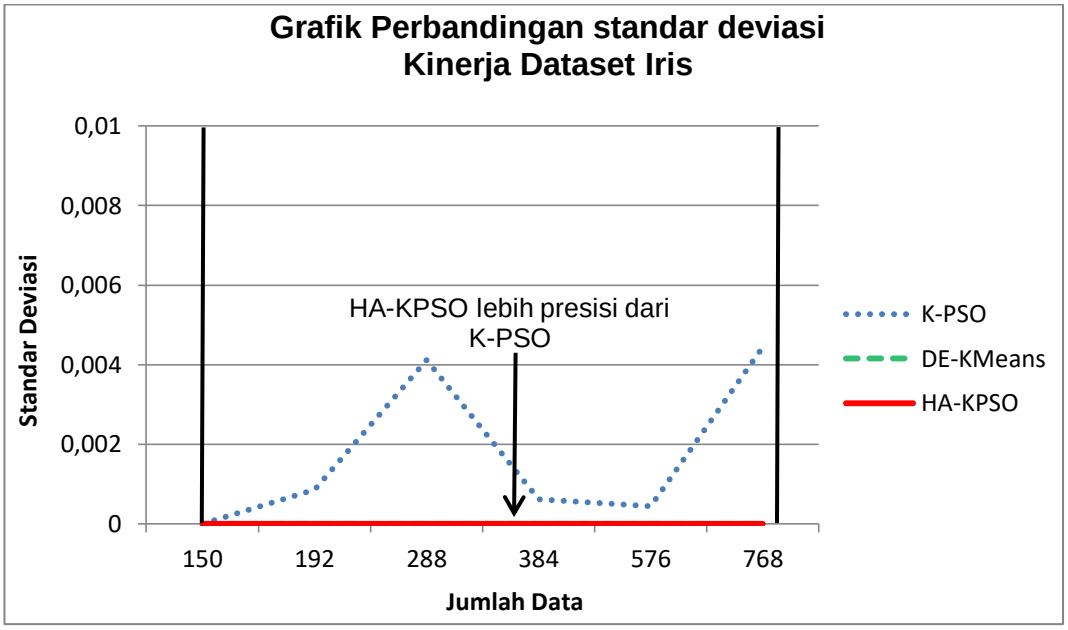
Tabel 4.14 menunjukkan rata-rata hasil pengujian DBI dan standar deviasi terhadap algoritma K-PSO, HA-KPSO dan DE-KMeans. Semakin kecil nilai DBI artinya semakin bagus kinerja algoritma. Sebaliknya semakin besar nilai DBI maka semakin buruk kinerja algoritma tersebut. Untuk memudahkan melihat perbandingan nilai DBI antar algoritma, maka hasil pengujian divisualisasikan ke dalam sebuah grafik yang ditampilkan di Gambar 4.26.



Gambar 4.26 Grafik kinerja DBI *Dataset* Iris

Berdasarkan grafik yang ditampilkan di Gambar 4.26 sumbu x merepresentasikan jumlah data dan sumbu y merepresentasikan nilai DBI. Untuk memudahkan membaca grafik maka nilai DBI pada Tabel 4.14 di*invers* terlebih dahulu sehingga semakin tinggi garis pada sumbu y artinya semakin baik kinerja algoritma. Sebaliknya semakin rendah garis pada sumbu y artinya semakin buruk kinerja algoritma. Dapat dilihat pada grafik bahwa algoritma HA-KPSO memiliki kinerja yang lebih dibandingkan dengan algoritma K-PSO untuk *dataset* iris. Hal ini disebabkan oleh perbedaan jumlah *cluster* yang digunakan dalam proses *clustering*. Cluster yang digunakan pada HA-KPSO lebih tepat sehingga hasil DBI nya pun lebih baik. Jika dibandingkan dengan Algoritma DE-KMeans, algoritma HA-KPSO mempunyai kinerja yang sama baiknya.

Selanjutnya, untuk melihat kepresisian algoritma maka dilakukan perbandingan terhadap standar deviasi dari rata-rata DBI yang dapat dilihat di Gambar 4.27.



Gambar 4.27 Grafik Standar Deviasi Kinerja *Dataset* Iris

Berdasarkan grafik yang ditampilkan di Gambar 4.27, dapat dilihat standar deviasi algoritma HA-KPSO untuk setiap jumlah data adalah 0, Artinya setiap 10 kali pengujian hasil clustering yang didapatkan selalu sama. Jika dibandingkan dengan algoritma K-PSO dan DE-KMeans, Algoritma HA-KPSO mempunyai standar deviasi yang lebih baik dari K-PSO dan Standar deviasi yang sama baiknya dengan DE-KMeans.

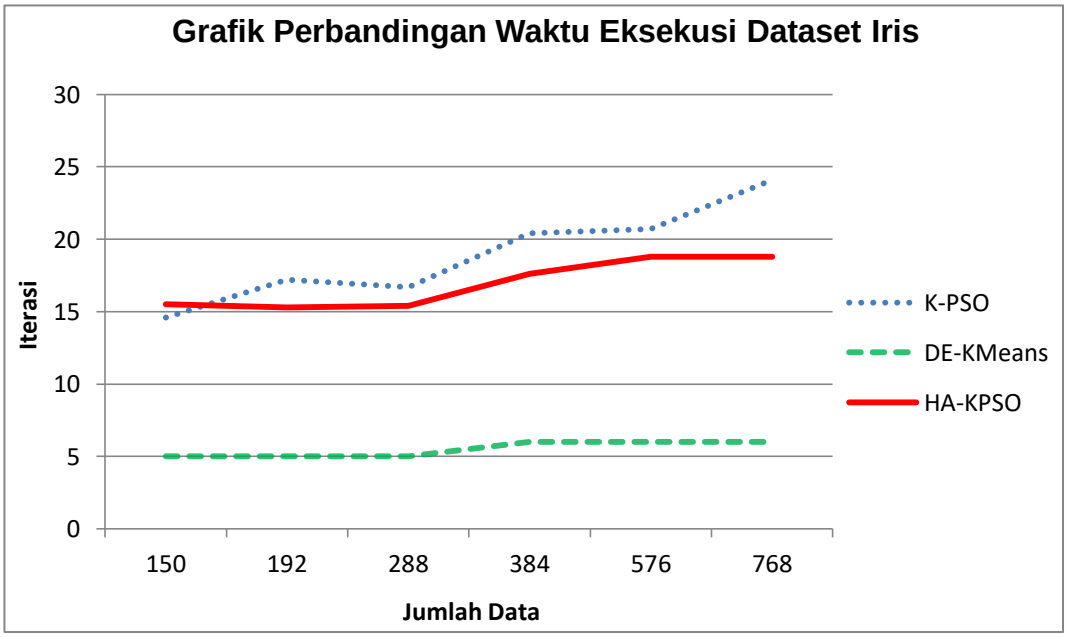
Untuk mengetahui *cost* yang dibutuhkan dalam menghasilkan hasil *clustering*, dilakukan pengujian terhadap waktu eksekusi algoritma dengan melihat jumlah perulangan (*iterasi*) dari proses clustering algoritma tersebut. Waktu eksekusi algoritma K-PSO, HA-KPSO dan DE-KMeans pada *dataset* iris disajikan di Tabel 4.15.

Tabel 4.15 Waktu Eksekusi dan Standar Deviasi *Dataset* Iris

Jumlah data	Waktu Eksekusi			Standar Deviasi		
	K-PSO	DE-KMeans	HA-KPSO	K-PSO	DE-KMeans	HA-KPSO
24	14,6	5	15,5	1,0749677	0	0,707106781
48	17,2	5	15,3	1,229272594	0	1,159501809
96	16,7	5	15,4	1,418136492	0	1,264911064
192	20,4	6	17,6	2,118699811	0	1,173787791

Jumlah data	Waktu Eksekusi			Standar Deviasi		
	K-PSO	DE-KMeans	HA-KPSO	K-PSO	DE-KMeans	HA-KPSO
384	20,7	6	18,8	1,82878223	0	1,988857852
768	24,1	6	18,8	2,024845673	0	2,097617696

Tabel 4.15 menunjukkan waktu eksekusi serta standar deviasi untuk dari algoritma K-PSO, HA-KPSO dan DE-KMeans. Semakin besar nilai iterasi artinya semakin lama waktu eksekusi algoritma dalam melakukan proses *clustering*. Sebaliknya semakin kecil nilai iterasi artinya semakin cepat waktu eksekusi algoritma dalam melakukan proses *clustering*. Untuk memudahkan melihat perbandingan waktu eksekusi dari ketiga algoritma tersebut maka hasil pengujian akan divisualisasikan ke dalam grafik yang ditampilkan di Gambar 4.28.

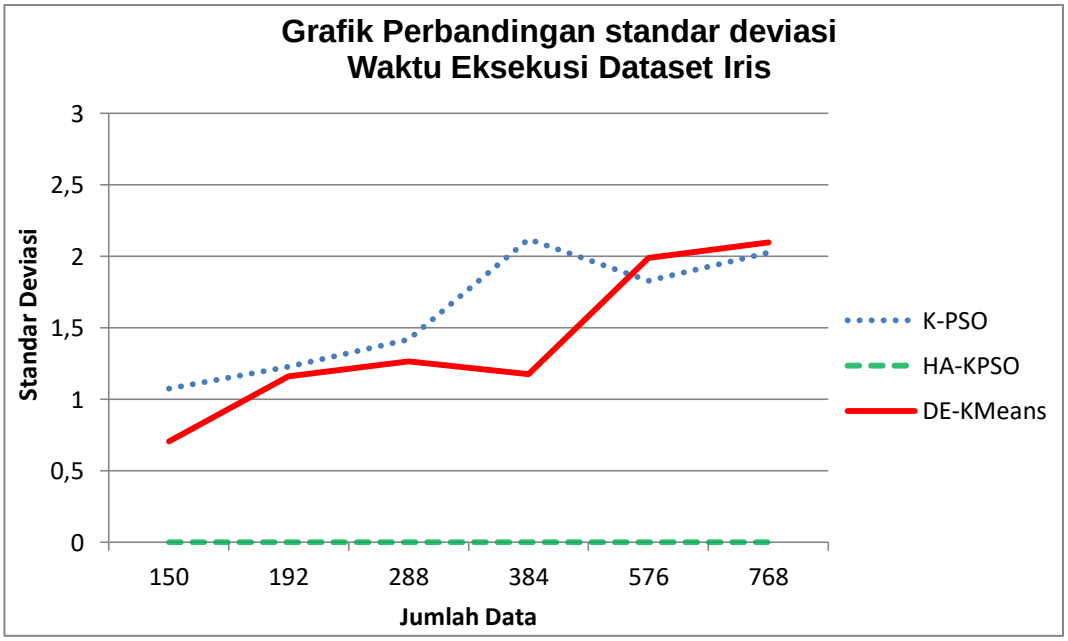


Gambar 4.28 Grafik Waktu Eksekusi *Dataset* Iris

Berdasarkan grafik yang ditampilkan di Gambar 4.28 sumbu x merepresentasikan jumlah data dan sumbu y merepresentasikan iterasi. Semakin tinggi garis pada sumbu y artinya semakin lama waktu eksekusi algoritma. Sebaliknya semakin rendah garis pada sumbu y artinya semakin cepat waktu eksekusi algoritma. Dapat dilihat bahwa algoritma HA-KPSO memiliki iterasi

yang lebih sedikit dibandingkan algoritma K-PSO. Sementara algoritma HA-KPSO memiliki jumlah iterasi yang lebih banyak dibandingkan dengan algoritma DE-KMeans.

Selanjutnya untuk melihat kepresisian algoritma maka dilakukan perbandingan terhadap standar deviasi iterasinya yang dapat dilihat di Gambar 4.29.



Gambar 4.29 Grafik Standar Deviasi Waktu Eksekusi *Dataset* Iris

Berdasarkan grafik yang ditampilkan di Gambar 4.29 dapat dilihat bahwa algoritma HA-KPSO mempunyai standar deviasi yang lebih buruk dibandingkan DE-KMeans yang selalu stabil di nilai angka 0. Jika dibandingkan dengan algoritma K-PSO, standar deviasi algoritma HA-KPSO lebih baik pada jumlah data ≤ 576. Namun saat jumlah data > 576, standar deviasi iterasi dari algoritma HA-KPSO menjadi lebih buruk dibandingkan algoritma K-PSO

4.1.4 Pengujian *Dataset* Diabetes

Pengujian *dataset* diabetes dilakukan sebanyak 10 kali menggunakan algoritma K-PSO, HA-KPSO dan DE-KMeans dengan variasi jumlah data yang dimulai dari 24 data s/d 768 data. Hasil pengujian yang akan ditampilkan berupa rata-rata nilai validasi *cluster*, standar deviasi rata-rata validasi *cluster*, rata-wata waktu eksesusi dan standar deviasi rata-rata waktu eksekusi.

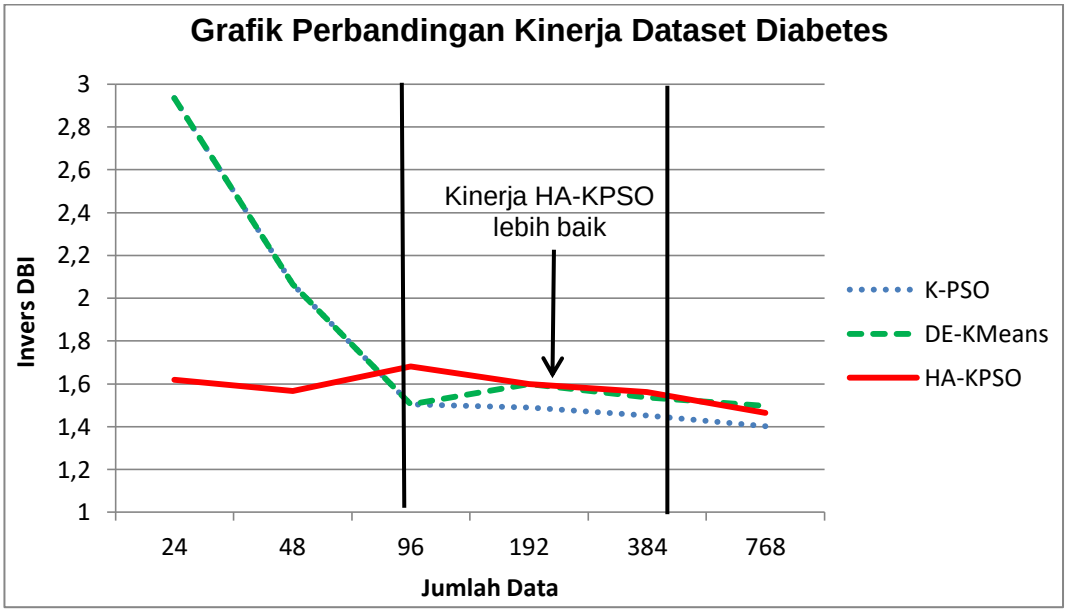
A. Hasil Pengujian Algoritma Bedasarkan nilai DBI

Pengujian hasil *clustering* mengunakan metode validasi *cluster* DBI pada algoritma K-PSO, HA-KPSO dan DE-KMeans untuk *dataset* diabetes dapat dilihat di Tabel 4.16.

Tabel 4.16 Hasil DBI dan Standar Deviasi Data Diabetes

Jumlah data	DBI			Standar Deviasi		
	K-PSO	DE-KMeans	HA-KPSO	K-PSO	DE-KMeans	HA-KPSO
24	0,340801194	0,3408012	0,61720281	0	0	0
48	0,484449178	0,4844492	0,638737279	0	0	0
96	0,664799092	0,6647991	0,594615676	0	0	0
192	0,671797071	0,6253485	0,625348554	0	0	0
384	0,68846577	0,6503302	0,630330167	0	0	0
768	0,713382373	0,6680979	0,668097944	0	0	0

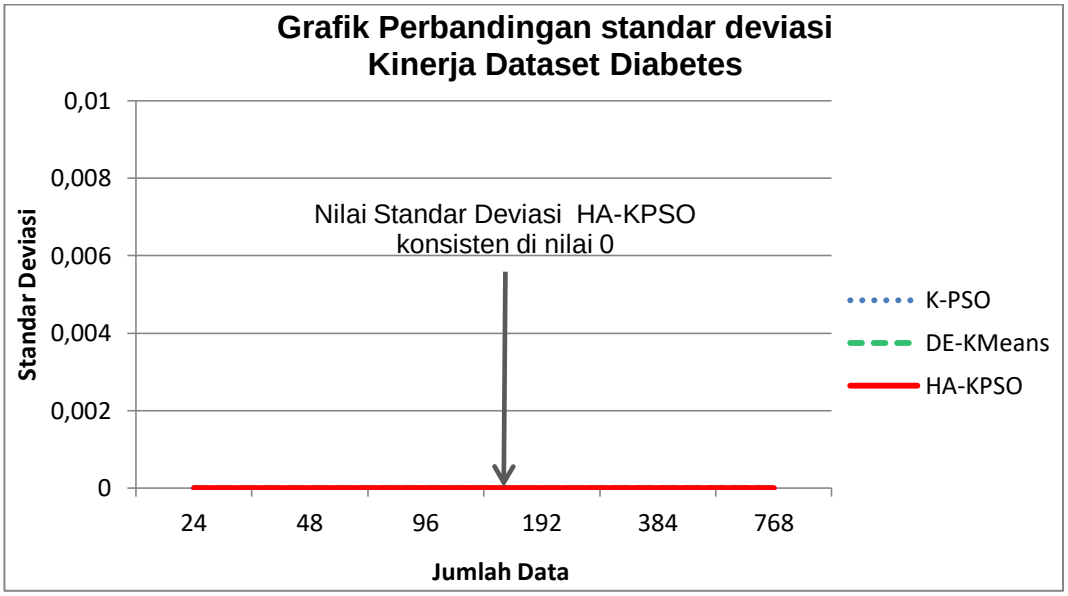
Tabel 4.16 menunjukkan rata-rata hasil pengujian DBI dan standar deviasi terhadap algoritma K-PSO, HA-KPSO dan DE-KMeans. Semakin kecil nilai DBI artinya semakin bagus kinerja algoritma. Sebaliknya semakin besar nilai DBI maka semakin buruk kinerja algoritma tersebut. Untuk memudahkan melihat perbandingan nilai DBI antar algoritma, maka hasil pengujian divisualisasikan ke dalam sebuah grafik yang ditampilkan di Gambar 4.30.



Gambar 4.30 Grafik kinerja DBI *Dataset* Diabetes

Berdasarkan grafik yang ditampilkan di Gambar 4.30 sumbu x merepresentasikan jumlah data dan sumbu y merepresentasikan nilai DBI. Untuk memudahkan membaca grafik maka nilai DBI pada Tabel 4.16 di*invers* terlebih dahulu sehingga semakin tinggi garis pada sumbu y artinya semakin baik kinerja algoritma. Sebaliknya semakin rendah garis pada sumbu y artinya semakin buruk kinerja algoritma. Dapat dilihat bahwa algoritma HA-KPSO memiliki kinerja yang lebih baik dibandingkan algoritma K-PSO pada kondisi jumlah data > 100 data. Sebaliknya pada jumlah data < 100 data, kinerja algoritma K-PSO lebih baik dari algoritma HA-KPSO. Jika Dibandingkan dengan algoritma DE-KMeans , algoritma HA-KPSO mempunyai kinerja yang lebih baik pada jumlah data di rentang 100 data s/d 500 data. Pada jumlah data < 100 dan > 500 data, algorimta DE-KMeans mempunyai kinerja yang lebih baik.

Selanjutnya, untuk melihat kepresisian algoritma maka dilakukan perbandingan terhadap standar deviasi dari rata-rata DBI yang dapat dilihat di Gambar 4.31.



Gambar 4.31 Grafik Standar Deviasi Kinerja *Dataset* Diabetes

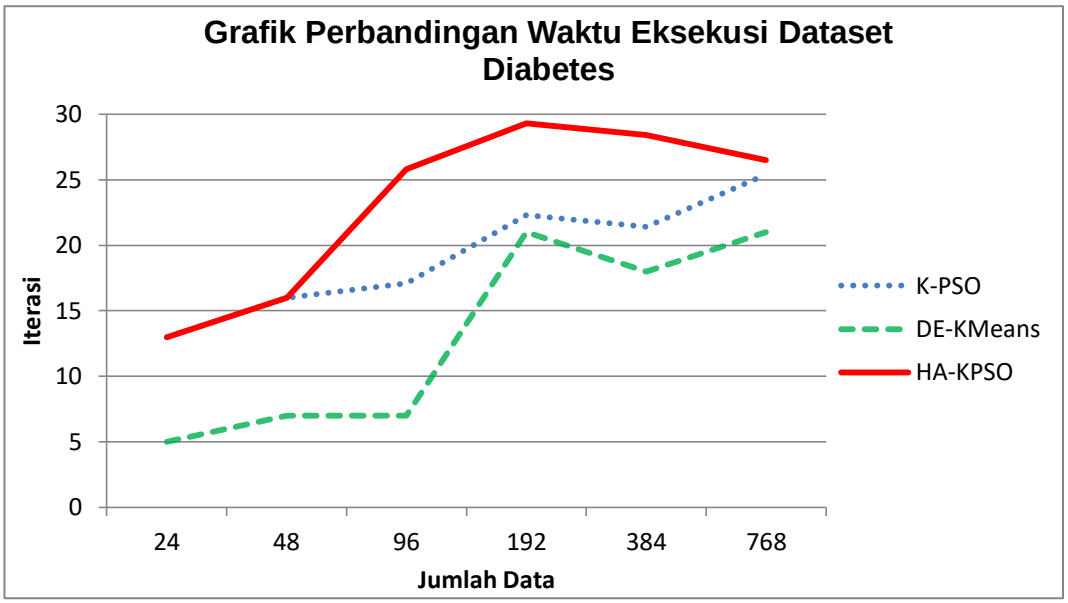
Berdasarkan grafik yang ditampilkan di Gambar 4.31 dapat dilihat standar deviasi algoritma HA-KPSO , KPSO, dan DE-Kmeans di setiap jumlah data yang diuji selalu bernilai 0, Artinya setiap dari 10 kali pengujian mendapatkan hasil clustering yang didapatkan selalu sama. Dapat disimpulkan pada *dataset* Diabetes ketiga algoritma mempunyai tingkat kepresisian yang sama.

Untuk mengetahui *cost* yang dibutuhkan dalam menghasilkan hasil *clustering*, dilakukan pengujian terhadap waktu eksekusi algoritma dengan melihat jumlah perulangan (*iterasi*) dari proses clustering algoritma tersebut. Waktu eksekusi algoritma K-PSO, HA-KPSO dan DE-KMeans pada *dataset* diabetes disajikan di Tabel 4.17.

Tabel 4.17 Waktu Eksekusi dan Standar Deviasi *Dataset* Diabetes

Jumlah data	Waktu Eksekusi			Standar Deviasi		
	K-PSO	DE-KMeans	HA-KPSO	K-PSO	DE-KMeans	HA-KPSO
24	13	5	13	0	0	0
48	16	7	16	0	0	0
96	17,1	7	25,8	0,875595036	0	2,043961296
192	22,3	21	29,3	1,636391694	0	2,668749187
384	21,4	18	28,4	1,837873167	0	1,646545205
768	25,4	21	26,5	2,118699811	0	1,416561177

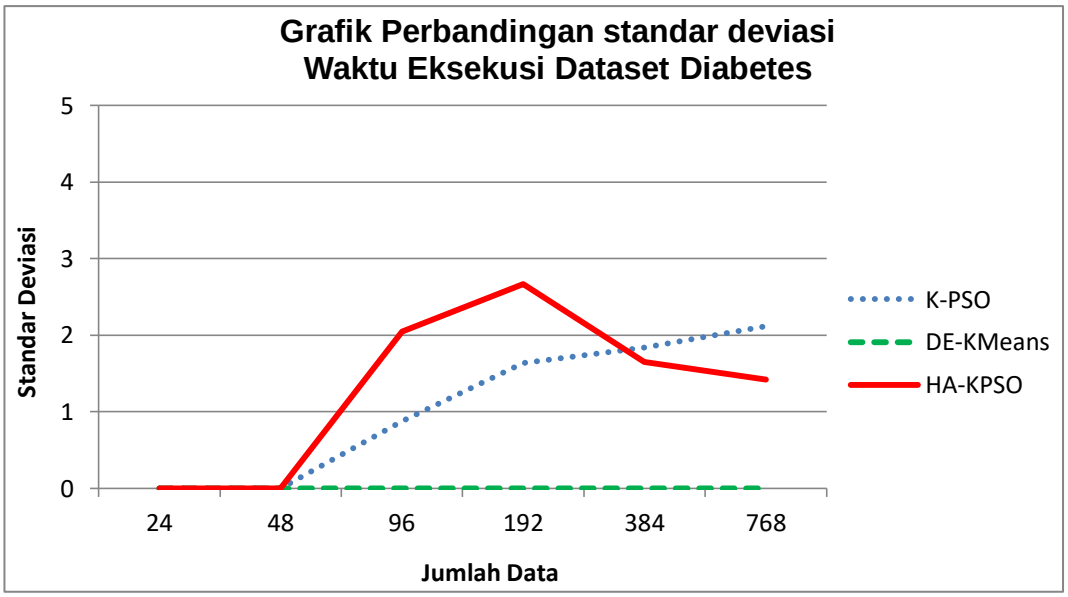
Tabel 4.17 menunjukkan waktu eksekusi serta standar deviasi untuk dari algoritma K-PSO, HA-KPSO dan DE-KMeans. Semakin besar nilai iterasi artinya semakin lama waktu eksekusi algoritma dalam melakukan proses *clustering*. Sebaliknya semakin kecil nilai iterasi artinya semakin cepat waktu eksekusi algoritma dalam melakukan proses *clustering*. Untuk memudahkan melihat perbandingan waktu eksekusi dari ketiga algoritma tersebut maka hasil pengujian akan divisualisasikan ke dalam grafik yang ditampilkan di Gambar 4.32.



Gambar 4.32 Grafik Waktu Eksekusi *Dataset* Diabetes

Berdasarkan grafik yang ditampilkan di Gambar 4.32 sumbu x merepresentasikan jumlah data dan sumbu y merepresentasikan iterasi. Semakin tinggi garis pada sumbu y artinya semakin lama waktu eksekusi algoritma. Sebaliknya semakin rendah garis pada sumbu y artinya semakin cepat waktu eksekusi algoritma. Dapat dilihat pada grafik bahwa algoritma HA-KPSO mempunyai iterasi yang lebih banyak dibandingkan algoritma K-PSO dan DE-KMeans di seluruh jumlah data yang diuji. Sedangkan DE-KMeans yang memiliki iterasi paling sedikit dari ketiga algoritma tersebut.

Selanjutnya, untuk melihat kepresisian algoritma maka dilakukan perbandingan standar deviasi iterasinya yang dapat dilihat di Gambar 4.33



Gambar 4.33 Grafik Standar Deviasi Waktu Eksekusi *Dataset* Diabetes

Berdasarkan grafik yang ditampilkan di Gambar 4.33 dapat dilihat bahwa algoritma HA-KPSO memiliki standar deviasi yang lebih buruk dibandingkan algoritma DE-KMeans. Jika dibandingkan dengan algoritma K-PSO, standar deviasi algoritma HA-KPSO lebih baik pada jumlah data > 384 data tetapi lebih buruk pada jumlah data < 384.

4.1.5 Pengujian *Dataset* Wine

Pengujian *dataset* wine dilakukan sebanyak 10 kali menggunakan algoritma K-PSO, HA-KPSO dan DE-KMeans dengan variasi jumlah data yang dimulai dari 150 data s/d 768 data. Hasil pengujian yang akan ditampilkan berupa rata-rata nilai validasi *cluster*, standar deviasi rata-rata validasi *cluster*, rata-wata waktu eksesusi dan standar deviasi rata-rata waktu eksekusi.

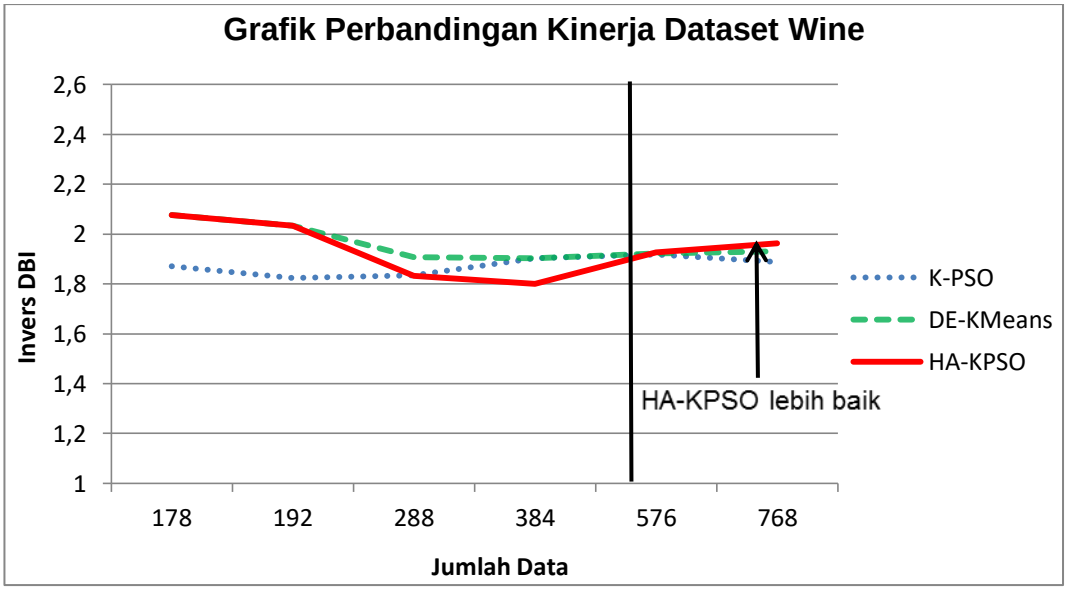
A. Hasil Pengujian Algoritma Bedasarkan nilai DBI

Pengujian hasil *clustering* mengunakan metode validasi *cluster* DBI pada algoritma K-PSO, HA-KPSO dan DE-KMeans untuk *dataset* wine dapat dilihat di Tabel 4.18.

Tabel 4.18 Hasil DBI dan Standar Deviasi *Dataset* Wine

Jumlah data	DBI			Standar Deviasi		
	K-PSO	DE-KMeans	HA-KPSO	K-PSO	DE-KMeans	HA-KPSO
178	0,5342432	0,481703	0,481703412	0	0	0
192	0,5482755	0,491930	0,49193081	0,00812170	0	0
288	0,5449379	0,545513	0,524184098	0,00763986	0,007405	0
384	0,5254713	0,555552	0,525785515	0,00296164	0	0
576	0,5212574	0,519242	0,520257479	0,00316227	0,002562	0
768	0,5364047	0,509196	0,518028724	0,00342347	0	0

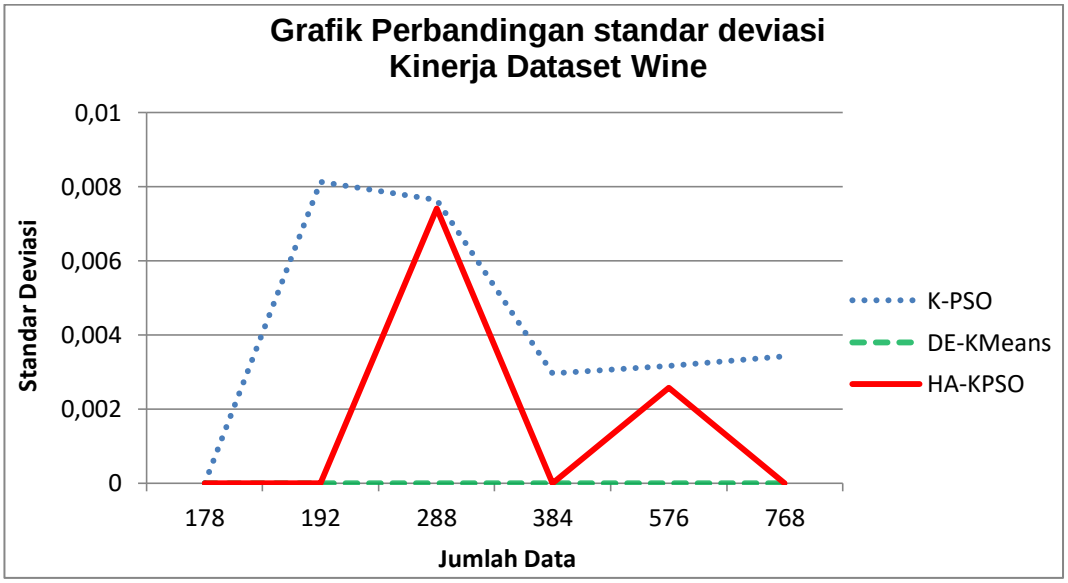
Tabel 4.18 menunjukkan rata-rata hasil pengujian DBI dan standar deviasi terhadap algoritma K-PSO, HA-KPSO dan DE-KMeans. Semakin kecil nilai DBI artinya semakin bagus kinerja algoritma. Sebaliknya semakin besar nilai DBI maka semakin buruk kinerja algoritma tersebut. Untuk memudahkan melihat perbandingan nilai DBI antar algoritma, maka hasil pengujian divisualisasikan ke dalam sebuah grafik yang ditampilkan di Gambar 4.34.



Gambar 4.34 Grafik kinerja DBI data Wine

Berdasarkan grafik yang ditampilkan di Gambar 4.34 sumbu x merepresentasikan jumlah data dan sumbu y merepresentasikan nilai DBI. Untuk memudahkan membaca grafik maka nilai DBI pada Tabel 4.18 di*invers* terlebih dahulu sehingga semakin tinggi garis pada sumbu y artinya semakin baik kinerja algoritma. Sebaliknya semakin rendah garis pada sumbu y artinya semakin buruk kinerja algoritma. Dapat dilihat pada grafik bahwa algoritma HA-KPSO memiliki kinerja yang lebih baik dibandingkan dengan algoritma DE-KMeans pada jumlah data > 500 data. Pada rentang jumlah data 200 s/d 500 data, Kinerja algoritma HA-KPSO lebih buruk dibandingkan algoritma K-PSO dan DE-KMeans. Pada jumlah data < 200 data, kinerja HA-KPSO sama baiknya dengan algoritma DE-KMeans dan lebih baik dibandingkan algoritma K-PSO.

Selanjutnya, untuk melihat kepresisian algoritma maka dilakukan perbandingan terhadap standar deviasi dari rata-rata DBI yang dapat dilihat di Gambar 4.35.



Gambar 4.35 Grafik Standar Deviasi Kinerja data Wine

Berdasarkan grafik yang ditampilkan di Gambar 4.35 dapat dilihat bahwa standar deviasi Algoritma HA-KPSO lebih baik dibandingkan algoritma K-PSO.

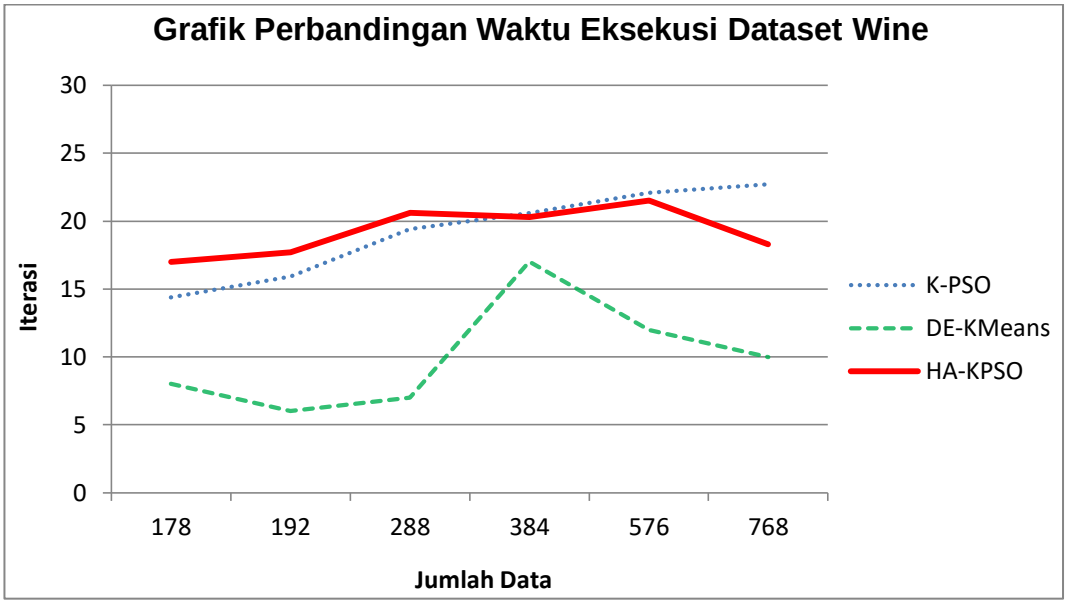
Jika dibandingkan dengan algoritma DE-Kmeans, algoritma HA-KPSO mempunyai standar deviasi yang lebih buruk.

Untuk mengetahui *cost* yang dibutuhkan dalam menghasilkan hasil *clustering*, dilakukan pengujian terhadap waktu eksekusi algoritma dengan melihat jumlah perulangan (*iterasi*) dari proses clustering algoritma tersebut. Waktu eksekusi algoritma K-PSO, HA-KPSO dan DE-KMeans pada *dataset* wine disajikan di Tabel 4.19.

Tabel 4.19 Waktu Eksekusi dan Standar Deviasi *Dataset* Wine

Jumlah data	Waktu Eksekusi			Standar Deviasi		
	K-PSO	DE-KMeans	HA-KPSO	K-PSO	DE-KMeans	HA-KPSO
24	14,4	8	17	1,173787791	0	1,333333333
48	15,9	6	17,7	1,100504935	0	1,159501809
96	19,4	7	20,6	1,837873167	0	2,319003617
192	20,6	17	20,3	1,577621275	0	1,418136492
384	22,1	12	21,5	2,024845673	0	1,779513042
768	22,7	10	18,3	1,82878223	0	1,766981104

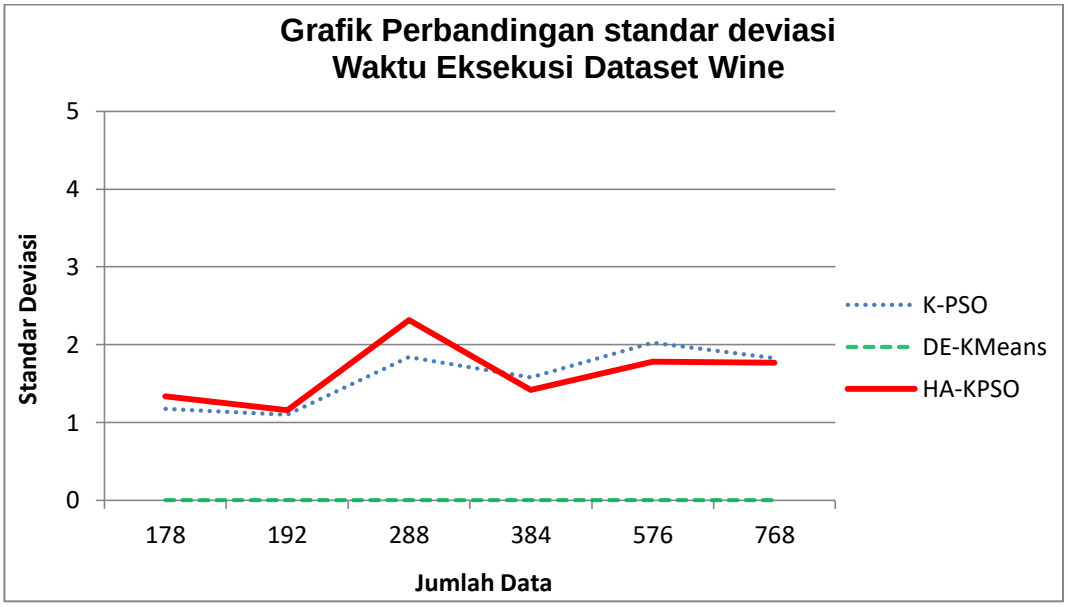
Tabel 4.19 menunjukkan waktu eksekusi serta standar deviasi untuk dari algoritma K-PSO, HA-KPSO dan DE-KMeans. Semakin besar nilai iterasi artinya semakin lama waktu eksekusi algoritma dalam melakukan proses *clustering*. Sebaliknya semakin kecil nilai iterasi artinya semakin cepat waktu eksekusi algoritma dalam melakukan proses *clustering*. Untuk memudahkan melihat perbandingan waktu eksekusi dari ketiga algoritma tersebut maka hasil pengujian akan divisualisasikan ke dalam grafik yang ditampilkan di Gambar 4.36.



Gambar 4.36 Grafik Waktu Eksekusi data Wine

Berdasarkan grafik yang ditampilkan di Gambar 4.36 sumbu x merepresentasikan jumlah data dan sumbu y merepresentasikan iterasi. Semakin tinggi garis pada sumbu y artinya semakin lama waktu eksekusi algoritma. Sebaliknya semakin rendah garis pada sumbu y artinya semakin cepat waktu eksekusi algoritma. Dapat dilihat pada grafik bahwa algoritma HA-KPSO memiliki iterasi yang lebih banyak dibandingkan algoritma K-PSO dan DE-KMeans pada jumlah data < 300 data. Pada jumlah data > 300 Algoritma HA-PSO mempunyai jumlah iterasi yang lebih sedikit dibandingkan algoritma K-PSO namun tetep lebih banyak dibandingkan algoritma DE-KMeans.

Selanjutnya, untuk melihat kepresisian algoritma maka dilakukan perbandingan terhadap standar deviasi iterasinya yang dapat dilihat di Gambar 4.37.



Gambar 4.37 Grafik Standar Deviasi Waktu Eksekusi data Wine

Berdasarkan grafik yang ditampilkan di Gambar 4.37 dapat dilihat bahwa algoritma HA-KPSO mempunyai standar deviasi yang lebih buruk dibandingkan K-PSO dan DE-KMeans pada jumlah data < 300 data. Pada jumlah data > 300 data, standar deviasi dari Algoritma HA-KPSO menjadi lebih baik dibandingkan algoritma K-PSO namun tetap lebih buruk dari algoritma DE-KMeans yang Nilai Standar Deviasinya konsisten di nilai 0.

4.1.6 Analisis Hasil Pengujian Algoritma Clustering terhadap Data Model

Bedasarkan Hasil pengujian data model yang di tampilkan di subbab 4.1, dapat diketahui bahwa algoritma HA-KPSO memiliki kinerja yang baik berdasarkan evaluasi hasil *clustering* menggunakan metode DBI pada jumlah data < 100 data dan jumlah data > 500 data untuk data dengan jumlah atribut ≥ 11. Hal ini dibuktikan dari membandingkan hasil *clustering* algoritma HA-KPSO dengan algoritma K-PSO dan DE-KMeans. HA-KPSO selalu lebih baik saat kondisi jumlah data lebih besar dari 500 untuk data dengan jumlah atribut 11 sampai 13. Pada pengujian *dataset* buatan dengan atribut ≤ 11, Kinerja Algoritma HA-KPSO lebih buruk dibandingkan DE-KMeans pada jumlah data > 500. Namun pada jumlah data < 100 kinerja HA-KPSO tetap lebih baik dibandingkan DE-KMeans

dan K-PSO. Bedasarkan analisis sebelumnya maka dapat disimpulkan algoritma HA-KPSO cocok untuk digunakan pada data yang mempunyai jumlah atribut 11 s/d 13 dan pada jumlah data < 100 dan > 500 data.

Bedasarkan Hasil pengujian data model yang di tampilkan di subbab 4.1, dapat diketahui bahwa algoritma HA-KPSO membutuhkan waktu eksekusi yang lama dalam melakukan proses *clustering*. Hal ini dibuktikan dari membandingkan hasil *clustering* algoritma HA-KPSO dengan algoritma K-PSO dan DE-KMeans. Algoritma HA-KPSO mempunyai jumlah iterasi yang lebih banyak di seluruh pengujian pada data model. Penyebab dari lamanya waktu eksekusi algoritma HA-KPSO adalah tahap proses pencarian *centroid* terbaik di algoritma tersebut. Algoritma HA-KPSO membutuhkan iterasi tambahan dalam pencarian solusi terbaik pada data. Selain itu bertambahnya jumlah data yang di proses juga dapat mempengaruhi lamanya waktu eksekusi yang dibutuhkan oleh algoritma HA-KPSO menjadi lebih lama. Bedasarkan analisis sebelumnya maka dapat disimpulkan bahwa waktu eksekusi proses clustering algoritma HA-KPSO lebih buruk dari DE-Kmeans dan K-PSO.

4.2 Hasil Pengujian Data Kasus

Pengujian terhadap data kasus dilakukan dengan melakukan implementasi algoritma HA-KPSO terhadap *dataset* IMDB. Kualitas dari hasil *clustering* yang terbentuk akan di ukur dengan metode *validasi cluster* DBI dan waktu dari proses *clustering* akan diukur menggunakan jumlah iterasi. Untuk mengetahui kinerja algoritma apakah dapat menghasilkan kualitas yang paling baik di data kasus, Maka algoritma HA-KPSO akan dibandingkan dengan hasil *clustering* dari algoritma HA-KPSO akan menggunakan algoritma K-PSO dan DE-KMeans yang diimplmentasikan pada data kasus yang sama.

4.2.1 Hasil *Clustering* Algoritma HA-KPSO Pada *Dataset* IMDB

Tabel 4.20 menampilkan hasil pengujian algoritma HA-KPSO bedasarkan nilai DBI, jumlah cluster dan jumlah iterasi terhadap dataset IMDB

Tabel 4.20 Hasil pengujian Algoritma HA-KPSO Pada Dataset IMDB

Jumlah *Cluster* Terbentuk	Nilai DBI	Jumlah Iterasi
6 Cluster	0,559321	31 Iterasi

Berdasarkan proses *clustering* menggunakan Algoritma HA-KPSO, diketahui bahwa pengelompokan *dataset* IMDB menggunakan algoritma HA-KPSO akan membentuk 6 kelompok *cluster*. Kinerja DBI yang dihasilkan dari evaluasi hasil *clustering* adalah 0,559321 dengan waktu eksekusi sebanyak 31 iterasi. Anggota pada masing-masing hasil *clustering* yang terbentuk dari algoritma HA-KPSO dapat dilihat di Tabel 4.21.

Tabel 4.21 Anggota Pada Setiap *Cluster* HA-KPSO

Cluster	Jumlah Anggota	Daftar Anggota
Cluster 1	91 Data	7, 11, 12, 18, 19, 20, 25, 34, 37, 38, 39, 40, 41, 46, 47, 48, 56, 57, 58, 60, 68, 82, 83, 86, 88, 89, 92, 95, 109, 114, 115, 116, 117, 118, 119, 120, 121, 125, 128, 133, 146, 149, 150, 153, 160, 169, 176, 180, 181, 182, 185, 186, 192, 202, 204, 205, 206, 213, 214, 215, 235, 239, 244, 249, 271, 275, 278, 279, 280, 292, 301, 317, 328, 336, 337, 344, 345, 347, 350, 358, 385, 387, 388, 389, 390, 391, 425, 427, 429, 465, 491
Cluster 2	87 Data	62, 84, 90, 106, 111, 141, 143, 161, 162, 168, 193, 199, 209, 226, 227, 237, 238, 252, 256, 260, 264, 268, 272, 273, 295, 303, 305, 307, 308, 309, 310, 311, 312, 313, 314, 316, 319, 326, 329, 333, 335, 340, 342, 373, 377, 378, 379, 381, 383, 392, 404, 405, 406, 409, 410, 411, 413, 414, 417, 419, 420, 423, 424, 464, 468, 470, 472, 473, 474, 475, 476, 477, 478, 479, 480, 481, 482, 483, 484, 485, 486, 487, 488, 489, 490, 499, 500
Cluster 3	113 Data	5, 14, 24, 28, 49, 51, 70, 76, 87, 102, 103, 104, 107, 131, 132, 134, 140, 142, 151, 157, 167, 171, 173, 175, 177, 179, 191, 194, 222, 223, 224, 225, 228, 229, 236, 241, 242, 243, 253, 257, 258, 259, 261, 265, 281, 284, 290, 291, 293, 294, 297, 298, 299, 302, 304, 306, 318, 320, 321, 323, 324, 325, 338, 339, 343, 346, 351, 359, 360, 361, 363, 364, 365, 367, 368, 369, 370, 371, 372, 374, 376, 380, 382, 396, 397, 398, 402, 407, 408, 412, 421, 422, 434, 449, 450, 451, 452, 453, 454, 455, 456, 457, 458, 460, 461, 462, 463, 466, 467, 469, 471, 495, 496
Cluster 4	19 Data	1, 3, 4, 8, 13, 17, 26, 27, 29, 32, 36, 43, 66, 123, 159, 183, 234, 428, 492
Cluster 5	47 Data	2, 6, 9, 10, 15, 21, 23, 30, 31, 33, 35, 45, 53, 54, 64, 67, 69, 78, 79, 80, 85, 94, 96, 97, 101, 105, 110, 112, 113, 124, 158, 172, 184, 198, 200, 201, 212, 230, 231, 232, 240, 263, 277, 330, 331, 332, 341
Cluster 6	143 Data	16, 22, 42, 44, 50, 52, 55, 59, 61, 63, 65, 71, 72, 73, 74, 75, 77, 81, 91, 93, 98, 99, 100, 108, 122, 126, 127, 129, 130, 135, 136, 137, 138, 139, 144, 145, 147, 148, 152, 154, 155, 156, 163, 164, 165, 166, 170, 174, 178, 187, 188, 189, 190, 195, 196, 197, 203,

		207, 208, 210, 211, 216, 217, 218, 219, 220, 221, 233, 245, 246, 247, 248, 250, 251, 254, 255, 262, 266, 267, 269, 270, 274, 276, 282, 283, 285, 286, 287, 288, 289, 296, 300, 315, 322, 327, 334, 348, 349, 352, 353, 354, 355, 356, 357, 362, 366, 375, 384, 386, 393, 394, 395, 399, 400, 401, 403, 415, 416, 418, 426, 430, 431, 432, 433, 435, 436, 437, 438, 439, 440, 441, 442, 443, 444, 445, 446, 447, 448, 459, 493, 494, 497, 498

Tabel 4.21 menampilkan jumlah *cluster* terbaik yang terbentuk dari algoritma HA-KPSO dan anggota pada masing-masing cluster yang diinisialisasikan dengan atribut id film dalam.

4.2.2 Perbandingan Hasil *Clustering* Pada *Dataset* IMDB

Untuk mengetahui apakah algoritma HA-KPSO merupakan algoritma yang paling tepat dalam menyelesaikan kasus yang mempunyai karakteristik data seperti *Dataset* IMDB, maka pada subbab ini akan dilakukan pembandingan terhadap hasil *clustering* beberapa algoritma lainnya. Sebagai algoritma pembanding, *dataset* IMDB diolah menggunakan algoritma K-PSO dan DE-KMeans kemudian dievaluasi menggunakan DBI. Hasil perbandingan nilai DBI algoritma HA-KPSO K-PSO, dan DE-KMeans dapat dilihat di Tabel 4.22

Tabel 4.22 Perbandingan Hasil *Clustering* pada *Dataset* IMDB

Algoritma	Jumlah *Cluster*	DBI	Waktu eksekusi
HA-KPSO	6 *Cluster*	0,559321	31 Iterasi
K-PSO	2 *Cluster*	0,566885	19 Iterasi
DE-KMeans	5 *Cluster*	0,561184	12 iterasi

Berdasarkan Tabel 4.22 terlihat bahwa algoritma HA-KPSO menghasilkan nilai DBI yang lebih kecil dibandingkan algoritma K-PSO dan DE-KMeans. Nilai DBI algoritma HA-KPSO 1,33 % lebih kecil dari nilai DBI algoritma K-PSO dan 0,33 % lebih kecil dari algoritma DE-KMeans. Hal tersebut menunjukkan bahwa kualitas hasil *clustering* dari algoritma HA-KPSO lebih baik daripada kualitas hasil *clustering* algoritma K-PSO dan DE-KMeans jika diukur berdasarkan nilai DBI.