

## PENGELOMPOKAN DATA HASIL TES KEPERIBADIAN 16PF SOPIR BUS MENGUNAKAN ALGORITMA GENETIKA

Luthfi Hidayat<sup>1</sup>, Wayan Firdaus Mahmudy<sup>2</sup>

Email: <sup>1</sup> mr.luthfihidayat@gmail.com, <sup>2</sup> wayanfm@ub.ac.id

(Naskah masuk: 8 Agustus 2016, diterima untuk diterbitkan: 8 September 2016)

### Abstrak

Tes kepribadian merupakan suatu metode tes yang disusun untuk mendeskripsikan bagaimana kecenderungan seseorang dalam bertindak laku maupun berpikir. Tes kepribadian sebenarnya hanya dapat dideskripsikan secara kualitatif karena sebenarnya kepribadian tidak dapat diukur. Namun, untuk membantu menjelaskan kepribadian seseorang dapat menggunakan bantuan angka sehingga hasil dari tes tersebut dapat di deskripsikan ke dalam bentuk kualitatif. Dalam penelitian sebelumnya, pengelompokan hasil data tes kepribadian 16PF dilakukan dengan metode *K – Means Clustering* yang digabung dengan metode *Silhouette Coefficient*. Penelitian tersebut memiliki hasil maksimum SI (*Silhouette Index*) hingga 0.4341. Dalam penelitian kali ini, metode yang dapat digunakan untuk mengelompokkan dan menghitung seluruh data serta atribut yang diperoleh menggunakan Algoritma Genetika. Tahapan untuk mengelompokkan data menggunakan metode yang sama seperti penelitian sebelumnya yaitu *K – Means Clustering* dan untuk menghitung *cluster* diperlukan representasi kromosom agar dapat membangkitkan nilai *Centroid* untuk perhitungan *Silhouette Coefficient*. Representasi kromosom yang digunakan adalah *real code genetic algorithm* dimana representasi tersebut dibangkitkan secara random dengan interval tertentu. Dari pengujian yang dilakukan, sistem mampu memberikan nilai SI terbaik pada jumlah populasi 40, jumlah generasi 15, kombinasi cr 0.7 dan mr 0.3. Algoritma genetika mampu memberikan solusi optimal dibandingkan dengan penelitian sebelumnya dimana dengan jumlah data yang sama menghasilkan nilai SI yang lebih baik.

**Kata Kunci:** Algoritma genetika, *Personality Factor*, *Clustering*, *K – Means*, *Silhouette Coefficient*.

### Abstract

#### Abstract

*Personality Test is a test method developed to describe how the tendency of a person's behavior and thinking. Actually, personality tests can only be described qualitatively because actual personality cannot be measured. However, figures can be used to help explaining an individual's personality, thus test results could be described into qualitative terms. In previous research, grouping data results 16PF personality test was conducted using K - Means Clustering combined with Silhouette Coefficient methods. The study has a maximum performance in terms of SI (silhouette index) of 0.4341. In the present study, the method can be used to classify, count all the data and attributes that are obtained using Genetic Algorithms. Stages for classifying data using the same method as previous research, that K - Means Clustering and to calculate cluster, required the representation of chromosomes in order to generate value of Centroid for the calculation Silhouette Coefficient. Chromosome representation used is real code genetic algorithm which is representations generated randomly with a certain interval. From the tests, systems are able to provide the best SI values in populations of 40, the number of generations 15, combination of cr mr are 0.7 and 0.3. Genetic algorithms are able to provide optimal solutions compared to a previous study in which the same amount of data to produce better value SI.*

**Keywords:** Genetic algorithms, *Personality Factor*, *Clustering*, *K – Means*, *Silhouette Coefficient*.

---

## 1. PENDAHULUAN

Berdasarkan hasil dari Sensus Penduduk 2010, jumlah penduduk yang ada di Indonesia mempengaruhi jumlah pertumbuhan kendaraan bermotor di Indonesia. Seperti yang telah dipublikasikan pada website Badan Pusat Statistik Indonesia, pada tahun 2013 tercatat jumlah unit mobil penumpang sebesar 11.484.514, bis 2.286.309, mobil barang 5.615.494 dan sepeda motor 84.732.652 (Anonymous, 2010).

Proses sertifikasi sopir dapat dilakukan dengan berbagai cara melalui tahapan seleksi dari berbagai aspek yang dibutuhkan, salah satunya adalah aspek perilaku atau kepribadian sopir (Armindya, 2015). Dengan menggunakan salah satu metode yang digunakan untuk merekrut pegawai, tes 16 PF merupakan salah satu metode yang dapat diberikan untuk mengetahui kepribadian sopir. Tes 16PF merupakan kumpulan pertanyaan yang memiliki struktur dasar kepribadian dari setiap orang yang akan menghasilkan skor dari masing – masing

faktor kepribadian dari seseorang yang mengambil tes tersebut (Boyle et al., 2008).

Beberapa penelitian yang berkaitan dengan metode k-means telah dilakukan oleh Prakoso (2015) dan Blas et al. (2012). Penelitian yang dilakukan oleh Prakoso (2015) mengelompokkan data hasil tes kepribadian dengan menggunakan metode k-means. Masalah yang muncul pada pengelompokkan data hasil tes tersebut adalah beberapa sopir yang lulus ataupun tidak lulus sertifikasi. Hasil dari penelitian tersebut adalah mengetahui kualitas sopir dari pengelompokkan data *raw score* sehingga pada tahapan pengujian didapat nilai *silhouette coefficient* hingga 0,4341 (Prakoso, 2015).

Penelitian berikutnya dilakukan oleh Blas et al. (2012). Peneliti mencoba untuk mengelompokkan Algoritma Genetika di dalam permasalahan klustering dengan pendekatan yang berbeda. Pengujian Algoritma Genetika dilakukan dengan membandingkan pendekatan pengelompokkan klasik seperti k-means dan DBSCAN sehingga mendapatkan hasil yang sangat baik (Blas et al., 2012). Dari hasil kedua penelitian tersebut dapat dikatakan bahwa, metode Algoritma Genetika cukup baik digunakan dalam pengelompokkan data hasil tes.

Berdasarkan uraian yang telah dijelaskan, penulis mengusulkan judul penelitian yang berjudul “Pengelompokkan Data Hasil Tes Kepribadian 16PF Sopir Bus Menggunakan Metode Algoritma Genetika”.

## 2. METODE

### a. 16 PF Questionnaire

Menurut R.B Cattell (Revelle, 2014), faktor pemeriksaan kepribadian menggunakan 16 *PF Questionnaire* dapat digunakan untuk anak – anak maupun orang dewasa. Riset yang dilakukan oleh Cattell menjelaskan bahwa kamus yang dibuat oleh Allport dan Odbert (1936) yang berisi 18000 daftar kata sifat yang mendeskripsikan kepribadian dapat diringkas menjadi 4500 dengan cara menggabungkan istilah yang mempunyai makna serupa. Kemudian Cattell melakukan penelitian serupa dengan mengkombinasikan istilah sebelumnya dan setelah melalui proses analisis faktor pada ciri kepribadian, dari 171 ciri kepribadian, Cattell dapat mengurangi lagi jumlah jenis kepribadian menjadi 31 ciri kepribadian kemudian dengan penelitian yang melibatkan beberapa alat ukur kepribadian dari beberapa sumber, Cattell akhirnya merumuskan jumlah jenis kepribadian menjadi 12 ciri dan 4 yang lainnya berdasarkan penelitian sebelumnya sehingga penelitian tersebut dinamakan 16 *Personality Factor Questionnaire* (Aiken & Groth, 2006).

### b. Data Uji

Sumber data yang digunakan dalam penelitian ini mengadopsi dari penelitian sebelumnya yang dilakukan oleh Ryo Cahyo P. (2015), yaitu data *raw score* yang merupakan data primer pada penelitian tersebut.

### c. Konsep Pengelompokan

Di dalam proses klasifikasi data dibutuhkan syarat yang harus dipenuhi, antara lain mengelompokkan masing-masing data yang telah diurutkan sesuai dengan karakteristik kelas dari data tersebut. Untuk itu, sebuah model yang akan digunakan harus memiliki konsep prediksi agar dapat mengelompokkan data yang baru sesuai dengan kelas dari data tersebut.

Pada proses pengelompokan dibutuhkan suatu cara untuk mengidentifikasi setiap data yang akan bergabung di sebuah kelompok agar data tersebut mirip satu sama persis dengan lainnya, cara tersebut dinamakan dengan Analisis Kelompok. Analisis Kelompok merupakan suatu proses yang mencakup banyak teknik beragam untuk menemukan struktur atau karakteristik di dalam objek yang diidentifikasi. Tujuan dari analisis kelompok adalah mengidentifikasi sebuah data atau objek yang telah bergabung dengan kelompok sesuai dengan karakteristik maupun kriteria yang telah diberi keterangan sebelumnya serta data yang berbeda dapat dikelompokkan sesuai dengan kategori masing-masing (Kumar, 2013).

### d. K-means clustering

*k-means clustering* adalah salah satu metode dalam pengelompokan data yang membagi sejumlah data ke dalam 2 kelompok atau lebih. Metode *k-means* memiliki tujuan agar dapat meminimalisir fungsi objek yang akan diatur dalam proses *clustering* pada umumnya sehingga metode tersebut dapat memaksimalkan variasi data antar *cluster* (Agusta, 2007).

Proses pengelompokan dengan menggunakan metode *k-means* memiliki beberapa tahapan sebagai berikut:

1. Menentukan jumlah *cluster*/kelompok
2. Mengalokasikan data ke dalam kelompok secara *random*.
3. Menghitung *centroid* dari data yang ada pada masing – masing *cluster*
4. Mengalokasikan data ke dalam *centroid* terdekat.
5. Mengulangi langkah ke-3 apabila masih ada data yang berpindah *cluster* namun pada nilai *centroid* berada pada diatas *threshold* yang ditentukan atau jika perubahan nilai pada fungsi objektif masih berada di atas nilai *threshold* dengan batasan iterasi

maksimum.

Dalam tahapan ke-3 dalam perhitungan *centroid*, setiap kelompok dihitung nilai *mean* (nilai rata – rata) pada seluruh nilai data di setiap fitur atributnya secara berkala. Jarak antara dua titik  $x_1$  dan  $x_2$  dihitung dengan menggunakan *euclidean distance* pada Persamaan (1) dan untuk fungsi objektif dihitung dengan menggunakan Persamaan (2) (Agusta, 2007).

$$D_{L_2}(x_2, x_1) = |x_2 - x_1|_2 = \sqrt{\sum_{j=1}^p (x_{2j} - x_{1j})^2} \quad (1)$$

$$F = \sum_{j=1}^N \sum_{i=1}^K a_{ji} D(x_j, C_i) \quad (2)$$

#### e. Silhouette Coefficient

*Silhouette Coefficient* merupakan suatu metode yang digunakan untuk melihat kualitas dan kekuatan *cluster* dimana pada *cluster* tersebut dapat diukur seberapa baik suatu objek ditempatkan dalam suatu cluster. Metode ini merupakan gabungan dari metode *cohesion* dan *separation* (Wahyuni dkk., 2016).. *Silhouette Coefficient* memiliki beberapa tahapan dalam perhitungannya yaitu sebagai berikut:

1. Menghitung rata – rata jarak  $a_i$  dan  $b_i$  dari data yang terpilih dengan menggunakan Persamaan (3) dan Persamaan (4).

$$a_i^j = \frac{1}{m_j - 1} \sum_{r=1}^{m_j} d(x_i^j, x_r^j) \quad (3)$$

$$i = 1, 2, \dots, m_j$$

$$b_i^j = \min \left\{ \frac{1}{m_n} \sum_{r=1}^{m_n} d(x_i^j, x_r^n) \right\}, i = 1, 2, \dots, m_n \quad (4)$$

Persamaan (3) merupakan persamaan untuk pencarian nilai  $a_i$  dimana nilai  $a_i$  adalah rata-rata jarak data ke- $i$  terhadap semua data lainnya dalam satu klaster.  $m_j$  merupakan jumlah data yang terdapat pada satu cluster.

Persamaan (4) adalah persamaan untuk mencari nilai  $b_i$  adalah hasil yang didapatkan melalui perhitungan rata-rata jarak  $a_i$  terhadap seluruh data dari cluster lain, kemudian dari hasil rata-rata dari  $b_i$  diambil perolehan nilai minimum yang akan digunakan untuk pencarian nilai  $S_i$ .

2. Setelah seluruh rata-rata jarak dari seluruh data selesai dicari, kemudian mencari nilai terkecil dari cluster yang telah dicari sebelumnya.

3. Nilai  $S_i$  dapat ditemukan dengan cara seperti pada Persamaan (5).

$$SI_i^j = \frac{b_i^j - a_i^j}{\max\{a_i^j, b_i^j\}} \quad (5)$$

Pada Persamaan (5) nilai yang digunakan dalam perhitungan  $b_i - a_i$  adalah nilai minimum dari kelompok yang diselisihkan dengan jarak rata-rata  $a_i$ . Perhitungan tersebut dilakukan agar mendapatkan nilai maksimum dari kedua kelompok. Apabila hasil dari seluruh *silhouette coefficient* tiap cluster telah ditemukan, perhitungan  $S_i$  global dapat dilakukan dengan cara menghitung rata-rata dengan membagi seluruh nilai  $S_i$  dengan sejumlah data yang ada.

#### f. Algoritma Evolusi

Algoritma evolusi merupakan inti dari bagian yang dapat menunjang perkembangan optimasi dari data yang sedang dijadikan objek penelitian. Algoritma evolusi memiliki beberapa subset yang dapat dijadikan acuan, seperti Algoritma genetika (*Genetic algorithm*), Pemrograman genetika (*Genetic programming*), Strategi evolusi (*Evolution strategies*), dan Pemrograman evolusi (*Evolutionary programming*) (Mahmudy, 2015). Dari beberapa subset yang telah disebutkan, Algoritma genetika adalah algoritma yang paling sering diterapkan di dalam kehidupan sehari-hari serta dalam penanganan permasalahan yang kompleks (Mahmudy, Marian & Luong, 2013). Struktur awal algoritma genetika berdasarkan pada representasi kromosom (individu) yang dipilih secara acak dengan susunan tertentu. Setiap individu tersebut akan digabungkan menjadi suatu kelompok yang sering disebut dengan populasi.

Seperti yang telah dijelaskan oleh Mahmudy (2015), setiap permasalahan diubah menjadi sebuah kromosom. Kromosom tersebut disusun berdasarkan kumpulan individu yang menggambarkan setiap variabel keputusan yang akan digunakan sebagai solusi. Pada tiap kromosom memiliki nilai *fitness* yang didapatkan melalui fungsi *fitness* yang akan digunakan sebagai penilaian kriteria baik atau tidaknya nilai kromosom yang akan dimasukkan ke dalam algoritma genetika. Setelah merumuskan kromosom tersebut, algoritma genetika akan menghasilkan hasil terbaik berdasarkan perulangan yang dilakukan oleh metode tersebut dan hasil dari perulangan tersebut akan menghasilkan hasil terbaik yang akan diuraikan (*decoding*) menjadi solusi mendekati optimum.

Dalam penelitian ini, perhitungan nilai *fitness* dapat dicari dengan menggunakan perhitungan *Silhouette Coefficient* dimana perhitungan tersebut digunakan untuk mencari nilai  $S_i$  yang

dianggap sebagai nilai *fitness*. Berikut merupakan persamaan perhitungan dalam mengetahui nilai *fitness*.

$$\alpha_i^j = \frac{1}{m_j - 1} \sum_{r=1}^{m_j} d(x_i^j, x_r^j) \quad (6)$$

$$i = 1, 2, \dots, m_j$$

$$b_i^j = \min \left\{ \frac{1}{m_n} \sum_{r=1}^{m_n} d(x_i^j, x_r^n) \right\}, i$$

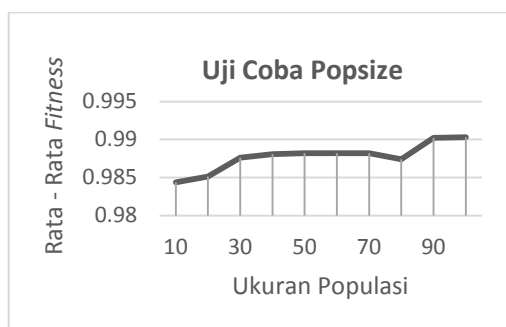
$$= 1, 2, \dots, m_n$$

$$Fitness = \frac{b_i^j - \alpha_i^j}{\max\{\alpha_i^j, b_i^j\}} \quad (8)$$

### 3. HASIL DAN PEMBAHASAN

#### a. Uji Coba Ukuran Populasi

Uji coba pertama kali dilakukan dengan cara uji coba ukuran populasi. Pada pengujian ini jumlah generasi yang digunakan adalah 5 dengan *crossover rate* sebesar 0,6 dan *mutation rate* sebesar 0,4. Jumlah data yang digunakan adalah 20 data. Setiap ukuran populasi dilakukan pengujian sebanyak 10 kali percobaan sehingga diperoleh nilai rata – rata *fitness*. Ukuran populasi diuji pada kelipatan 10, mulai dari 10 hingga 100, sehingga terdapat 10 ukuran popsize yang berbeda. Hasil grafik uji coba ukuran populasi dapat dilihat pada Gambar 1.



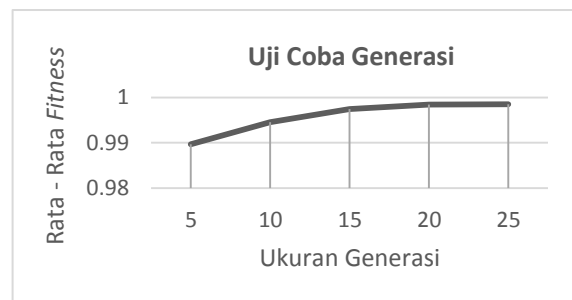
Gambar 1 Hasil Uji Coba Populasi

Titik optimal terdapat pada ukuran populasi 40 dengan nilai rata-rata *fitness* 0.988077. Sedangkan nilai *fitness* terendah terdapat pada ukuran populasi 10 dengan nilai rata-rata *fitness* 0.984378. Ukuran populasi yang besar belum tentu dapat menghasilkan solusi yang lebih baik. Hal ini dapat dilihat dari menurunnya nilai rata-rata *fitness* yang juga mencapai titik stabil pada ukuran populasi 80.

#### b. Uji Coba Banyaknya Generasi

Uji coba jumlah generasi dilakukan setelah pengujian ukuran populasi dilakukan agar mendapatkan hasil perubahan nilai *fitness* yang

konvergen. Dalam uji coba jumlah generasi, penggunaan parameter generasi yang semakin banyak akan mengalami peningkatan perbaikan nilai rata – rata *fitness* sehingga nilai tersebut menjadi stabil atau konvergen hingga kelipatan jumlah generasi selanjutnya. Setiap generasi dilakukan pengujian nilai *fitness* sebanyak 10 kali, pengujian dilakukan dengan menggunakan parameter populasi optimal sebelumnya yaitu 40, nilai *crossover rate* 0.6, nilai *mutation rate* 0.4 dan jumlah data yang digunakan adalah 20. Hasil grafik uji coba jumlah generasi dapat dilihat pada Gambar 2.

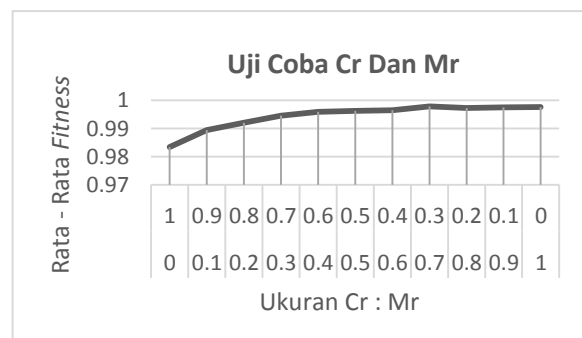


Gambar 2 Hasil Uji Coba Generasi

Titik optimal terdapat pada ukuran generasi 15 dengan nilai rata-rata *fitness* 0.997446. Sedangkan nilai *fitness* terendah terdapat pada ukuran generasi 5 dengan nilai rata-rata *fitness* 0.989715. Pengujian dihentikan pada ukuran generasi 25 karena tidak terjadi perubahan nilai yang signifikan.

#### c. Uji Coba Kombinasi Cr dan Mr

Uji coba kombinasi cr dan mr dilakukan setelah pengujian ukuran populasi dan pengujian jumlah generasi. Setiap kombinasi cr dan mr diuji coba sebanyak 10 kali. Nilai rata-rata tertinggi menunjukkan solusi yang paling mendekati optimal. Hasil grafik kombinasi cr dan mr ditunjukkan pada Gambar 3.



Gambar 3 Hasil Uji Coba Cr dan Mr

Titik optimal terdapat pada kombinasi cr dan mr 0.7 dan 0.3 dengan nilai rata-rata *fitness* 0.997768. Sedangkan nilai *fitness* terendah terdapat pada kombinasi 0 dan 1 dengan nilai rata – rata *fitness* 0.983357.

#### d. Uji Coba Parameter Terbaik

Berdasarkan pengujian parameter yang telah dilakukan sebelumnya, terdapat tiga pengujian parameter untuk mendapatkan nilai parameter terbaik yaitu ukuran populasi, banyak generasi serta kombinasi Cr dan Mr. Pada ketiga uji coba tersebut terdapat hasil optimal yang menunjukkan nilai *fitness* terbaik. Parameter yang digunakan untuk mendapatkan nilai optimal sebagai berikut.

- Popsize = 40
- Jumlah generasi = 15
- cr = 0.7
- mr = 0.3

Hasil dari pengujian dengan menggunakan nilai parameter terbaik dapat dilihat pada Tabel 1.

**Tabel 1 Hasil Pengujian Parameter Terbaik**

Nilai <i>fitness</i> percobaan ke-										Rata-rata <i>Fitne ss</i>
1	2	3	4	5	6	7	8	9	10	
0.997	0.997	0.998	0.996	0.998	0.997	0.998	0.998	0.998	0.998	0.99
532	847	955	617	471	505	054	083	107	025	792

Berdasarkan hasil pengujian yang ditunjukkan pada Tabel 1, tabel tersebut menunjukkan hasil rata – rata nilai *fitness* tertinggi yang didapatkan sebesar 0.99792.

### 3. SIMPULAN

Dalam penelitian ini, metode algoritma genetika berhasil diterapkan dalam proses pengelompokan data hasil tes kepribadian 16 PF sopir bus. Dari pengujian yang telah dilakukan maka didapatkan nilai parameter dalam algoritma genetika yang mampu memberikan hasil atau solusi yang dianggap optimal. Pada implementasi pengelompokan data hasil tes kepribadian 16PF sopir bus menggunakan algoritma genetika, solusi ukuran populasi terbaik adalah 40 popsize, solusi banyaknya generasi terbaik adalah 15, solusi kombinasi nilai *crossover rate* (cr) 0.7 dan *mutation rate* (mr) 0.3.

Pengelompokan dengan menggunakan algoritma genetika memberikan solusi yang lebih baik daripada pengelompokan dengan menggunakan *k – means clustering*. Berdasarkan nilai *Silhouette Index* (SI) yang dihasilkan pada penelitian sebelumnya, nilai SI yang dihasilkan dengan metode *k – means clustering* menggunakan *Silhouette Coefficient* tertinggi hingga 0,4341. Sedangkan nilai SI yang merupakan nilai *fitness* dari algoritma genetika menghasilkan nilai tertinggi hingga 0.99792.

Pada penelitian selanjutnya diharapkan dapat menggunakan data yang lebih bervariasi karena

semakin banyak variasi data maka akan semakin banyak pula variasi pengujian agar mendapatkan hasil yang lebih optimal. Hibridisasi dengan algoritma lain juga bisa dilakukan agar didapatkan hasil yang lebih baik (Mahmudy, 2014).

### 4. DAFTAR PUSTAKA

- Agusta, Y. 2007. “K-Means – Penerapan, Permasalahan Dan Metode Terkait.” STMIK STIKOM BALI, Denpasar, Bali 3.
- Aiken, L. R, & G. Groth. 2006. Psychological Testing and Assessment. Twelfth. Indeks.
- Anonymous. 2010. “Perkembangan Jumlah Kendaraan Bermotor Menurut Jenis Tahun 1987-2012.” [http://www.bps.go.id/website/tabelExcelIndo/indo\\_17\\_12.xls](http://www.bps.go.id/website/tabelExcelIndo/indo_17_12.xls). [Diakses pada tanggal 11 September 2015].
- Armindya, Y. R. 2015. “Ahok: Sertifikasi Sopir Transjakarta Wajib Hukumnya.” <http://metro.tempo.co/read/news/2015/06/25/083678174/ahok-sertifikasi-sopir-transjakarta-wajib-hukumnya>. [Diakses pada tanggal 12 September 2015].
- Blas, L.E. Agustín, Salcedo S. Sanz, S. Jimenez Fernandez, L. Carro Calvo, J. Del Ser, & J. A. Portilla Figueras. 2012. “A New Grouping Genetic Algorithm for Clustering Problems.” Journal Department of Signal Theory and Communications, Universidad de Alcalá, Madrid, Spain.
- Boyle, G. J., G. Matthews, & DH. Saklofske. 2008. “The SAGE Handbook of Personality Theory and Assessment: Personality Measurement and Testing.” New York: Sage Publications 2.
- Kumar, A. 2013. “Cluster Analysis.” Journal of I.A.S.R.I. Library Avenue, New Delhi.
- Mahmudy, W. F. 2015. *Dasar-Dasar Algoritma Evolusi*. Program Teknologi Informasi dan Ilmu Komputer, Universitas Brawijaya.
- Mahmudy, WF, Marian, RM & Luong, LHS 2013, 'Modeling and optimization of part type selection and loading problems in flexible manufacturing system using real coded genetic algorithms', *International Journal of Electrical, Computer, Electronics and Communication Engineering*, vol. 7, no. 4, pp. 251-260.
- Mahmudy, W. F. 2014, *Optimisation of Integrated Multi-Period Production Planning and Scheduling Problems in*

- Flexible Manufacturing Systems (FMS) Using Hybrid Genetic Algorithms*, School of Engineering, University of South Australia.
- Prakoso, Ryo C. 2015. "Pengelompokan Data Hasil Tes Kepribadian 16PF Sopir Bus Menggunakan Metode K-Means." DORO: Repository Jurnal Mahasiswa PTIIK. Fakultas Ilmu Komputer. Universitas Brawijaya. Malang.
- Revelle, W. 2014. "Raymond B. Cattell." Publication of The Personality Project.
- Wahyuni, I., Auliya, Y. A., Rahmi, A., Mahmudy, W. F. (2016). "Clustering Nasabah Bank Berdasarkan Tingkat Likuiditas Menggunakan Hybrid Particle Swarm Optimization dengan K-Means." *Jurnal Ilmiah Teknologi dan Informasi ASIA (JITIKA)* 10(2): 24-33.
- Trayasiwi, G. P. 2015. "Penerapan Metode Klastering Dengan Algoritma K-Means Untuk Prediksi Kelulusan Mahasiswa Pada Program Studi Teknik Informatika Strata Satu." Fakultas Ilmu Komputer, Universitas Dian Nuswantoro, Semarang.