

**Pembentukan Grup Homogen pada Pembelajaran Kolaboratif menggunakan *Performance Factor Analysis***  
***Group Formation on Collaborative Learning using Performance Factor Analysis***

Josua Deston Girsang<sup>1</sup>  
(1103121273)

Dade Nurjanah, Ir., M.T., Ph.D<sup>2</sup>  
(97730158-1)

Fakultas Teknik Informatika Universitas Telkom Bandung

joshuadeston@gmail.com<sup>1</sup>

dadnurjanah@gmail.com<sup>2</sup>

---

---

**Abstrak**

*Computer-Supported Collaborative Learning (CSCL)* adalah bentuk pembelajaran kolaboratif yang dibantu oleh komputer. CSCL memanfaatkan teknologi komputer dalam berbagai aspek, mulai dari untuk meningkatkan interaksi antar anggota kelompok, berbagi pengetahuan, hingga membagi dan menggabungkan pekerjaan kelompok tersebut. CSCL tidak dapat berjalan tanpa adanya *Computer-Supported Group Formation (CSGF)*.

Pada tugas akhir ini penulis menganalisis dan merancang CSGF yang cocok untuk pengimplementasian algoritma *Performance Factor Analysis*. Pengimplementasian dan pengujian sistem serta parameter – parameter yang dibutuhkan oleh *Performance Factor Analysis* ditentukan dengan pendekatan *bounded*. Sistem CSGF yang dibangun menggunakan *Stereotype model* yang menggunakan *knowledge base* dan *average performance* sebagai atribut yang menjadi *stereotype*. *Stereotype model* dibangun menggunakan *K-means clustering*. Grup yang dibentuk bersifat *ability group* dan beranggotakan 3 orang.

Parameter – parameter *Performance Factor Analysis* yaitu  $\beta$ ,  $\gamma$ , dan  $\rho$  berpengaruh terhadap homogenitas grup yang dibentuk. Ketiga parameter berbanding terbalik dan memiliki pengaruh yang tidak begitu besar dengan homogenitas grup yang dibentuk. Sebaliknya, inisialisasi *centroid* pada saat pembuatan *stereotype model* menggunakan *K-means* memiliki pengaruh yang signifikan terhadap grup yang dibentuk, termasuk homogenitasnya. Dimana beberapa inisialisasi *centroid* yang dilakukan menghasilkan grup yang homogenitasnya baik, dan beberapa tidak.

**Kata Kunci:** *stereotype model, user model, performance factor analysis, K-means clustering, knowledge base, average performance, homogenitas, grup*

---

**Abstract**

*Computer-Supported Collaborative Learning (CSCL)* is a form of Collaborative learning using Computer as a support. CSCL utilize computer-based technology on many aspects, whether to improve interaction, share knowledge, or break down tasks within the group members. CSCL can not be made without Computer-Supported Group Formation.

In this research, the author analyze and design a suitable CSGF with the implementation of a Performance Factor Analysis algorithm. Implementation, system testing, and determining parameters are all based on bounded approach. The CSGF system is consisted of a stereotype model which use student's knowledge base and average performance as stereotypes. Stereotype model is constructed using K-means clustering. The group which formed has the feature of an ability group and consist of 3 person for each group.

Performance Factor Analysis parameters, which is  $\beta$ ,  $\gamma$ , and  $\rho$  respectively, has an influence towards the homogeneity of the constructed groups. These parameters is inversely proportional and has little influence compared to the centroid initialization done in the K-means process. The initialization could make a significant change in the resulted group, included its homogeneity. Whereas some initialization produced a good homogeneity, but some others do not.

**Keyword:** *stereotype model, user model, performance factor analysis, K-means clustering, knowledge base, average performance, group*

---

**1. Pendahuluan**

**1.1 Latar Belakang**

Dalam beberapa tahun terakhir, penelitian tentang Intelligent Tutoring System (ITS) mendapat banyak kritik karena mengesampingkan aspek sosial dari proses belajar dan berkembang kearah pembelajaran individu yang membagikan pengetahuan lewat potongan-potongan kecil [2]. Sehingga mulai banyak penelitian yang berfokus pada pembelajaran kolaboratif, dimana pengembangannya didasari oleh teori yang

didefinisikan Vygotsky, yaitu bahwa belajar adalah sebuah proses kolaboratif dimana siswa mendapatkan pengetahuan dari interaksi sosial yang dialaminya [3].

Computer-Supported Collaborative Learning (CSCL) adalah bentuk pembelajaran kolaboratif yang dibantu oleh komputer. CSCL memanfaatkan teknologi komputer dalam berbagai aspek, mulai dari untuk meningkatkan interaksi antar anggota kelompok, berbagi pengetahuan, hingga membagi dan menggabungkan pekerjaan kelompok tersebut.

CSCL juga, dalam konteks adaptif, melakukan pembagian kelompok sesuai dengan pengelompokan yang diinginkan dan juga mengadaptasi pembelajaran yang diberikan kepada kelompok sesuai dengan performansi yang dilakukan oleh kelompok tersebut.

Sampai saat ini belum ada sistem adaptif dari CSCL yang cukup baik dikarenakan banyaknya faktor yang harus diperhitungkan dalam proses pembelajaran kolaboratif. Salah satu faktor yang cukup mempengaruhi adalah metode yang digunakan dalam pembentukan grup belajar.

Secara garis besar, pembentukan grup belajar didasari oleh dua jenis pengelompokan, yaitu pengelompokan homogen yang berdasarkan kesamaan karakteristik siswa yang tidak sering berubah, dan pengelompokan heterogen yang berdasarkan profil siswa seperti umur, jenis kelamin, etnis, maupun faktor-faktor lain. Pada pengelompokan homogen, salah satu karakteristik siswa yang dapat dijadikan acuan dalam membentuk grup adalah nilai atau performansi siswa dalam menjalani sebuah tes. Performansi yang dihasilkan akan sangat beragam untuk setiap siswa dan akan memiliki dampak yang berbeda untuk tiap jenis pengelompokan yang dilakukan. Oleh karena itu, dibutuhkan suatu sistem yang mengelompokkan siswa dengan grup yang sesuai.

Terdapat beberapa metode untuk memroses nilai menjadi sebuah ukuran performansi agar dapat dijadikan sebagai acuan dalam pembentukan grup, beberapa diantaranya adalah Knowledge Tracing, Learning Factor Analysis, dan Performance Factor Analysis. Ketiga metode tersebut sama – sama mengukur kemampuan siswa dalam menjalani sebuah tes dengan cara yang berbeda, namun Knowledge Tracing dan Learning Factor Analysis tidak begitu efisien dibanding Performance Factor Analysis [1][4]. Performance factor analysis juga langsung mengukur nilai performansi murni siswa tersebut tanpa mempertimbangkan faktor – faktor lain. Sehingga dengan menggunakan Performance Factor Analysis, dapat dicari average performance tiap siswa yang akurat dan menghasilkan grup homogen yang sesuai dengan mempertimbangkan average performance tiap siswa yang menjadi anggota grup yang telah dibentuk.

## 1.2 Perumusan Masalah

Ada 4 pertanyaan untuk dijawab pada tugas akhir ini, yaitu :

1. Bagaimana arsitektur sistem CSGF yang tepat untuk dapat mengimplementasikan PFA ke dalam Group Formation-nya ?
2. Bagaimana cara untuk menerjemahkan output dari PFA untuk menjadi acuan dalam pembentukan grup ?
3. Apa pengaruh parameter PFA terhadap homogenitas grup yang dibentuk ?

4. Bagaimana komposisi parameter yang sesuai agar dihasilkan sebuah pembentukan grup yang benar – benar homogen ?

Adapun batasan masalah dari tugas akhir ini adalah :

1. Parameter – parameter dari formula Performance Factor Analysis yaitu  $\beta$ ,  $\gamma$ , dan  $\rho$  ditentukan sendiri nilainya sesuai dengan batasan yaitu 0 – 0.5, 0 – 0.5, dan 0 – 0.1 untuk setiap parameter tersebut [1][7].
2. Siswa yang akan dibentuk kelompoknya merupakan mahasiswa Fakultas Informatika Universitas Telkom yang mengambil mata kuliah dan praktikum Dasar Algoritma dan Pemrograman di tahun ajaran 2015-2016.

## 1.3 Tujuan

Tujuan dari tugas akhir ini adalah:

1. Mendesain skenario yang cocok untuk pembagian grup yang menjadikan average performance sebagai acuan.
2. Membentuk grup yang homogen dan menyisakan siswa tanpa kelompok seminimal mungkin.
3. Menganalisis efek dari perubahan parameter pada PFA terhadap homogenitas grup.

## 2. Dasar Teori

### 2.1. Collaborative Learning

Collaborative Learning atau pembelajaran kolaboratif secara luas dapat didefinisikan sebagai salah satu pendekatan dalam belajar yang melibatkan kerjasama secara intelektual diantara siswa dan siswa, ataupun siswa dan pengajar [10]. Pembelajaran kolaboratif sendiri bertujuan untuk membuat suasana belajar semirip mungkin dengan pembelajaran pada dunia nyata. Karena menurut Vygotsky, pada hakekatnya belajar adalah sebuah proses kolaboratif dimana siswa mendapatkan pengetahuan dari interaksi sosial yang dialaminya [3].

Jeff Golub mengemukakan bahwa keunggulan dari Collaborative Learning adalah dari fitur utama pada sistemnya yang mendukung siswa untuk berkomunikasi : “Karena siswa memang seharusnya saling berkomunikasi satu sama lain... dan dari komunikasi inilah sebagian besar pembelajaran terjadi” (Golub, 1988) [11].

Selain itu, ketika kita menggunakan Collaborative Learning sebagai metode belajar karena kita meyakini bahwa metode tersebut membantu siswa untuk belajar secara efektif, Collaborative Learning juga secara tidak langsung mengembangkan social skill siswa seperti tata cara berkomunikasi, toleransi, dan kerjasama, dimana banyak pengajar yang memiliki pandangan bahwa belajar adalah sebuah proses yang kompleks dan jauh dari sekedar menguasai konten materi yang diberikan [10].

Dari teori yang telah dikemukakan tersebut dapat disimpulkan bahwa tujuan sebenarnya dari pembelajaran kolaboratif adalah agar siswa yang terlibat dalam proses pembelajaran dapat berkembang dan memperoleh pengetahuan baru dari interaksi yang dilakukan siswa tersebut selama belajar.

## 2.2. Computer-Supported Group Formation

Setiap sistem Computer-Supported Collaborative Learning atau CSCL memiliki cara masing-masing dalam membentuk grup untuk proses pembelajaran yang akan dilakukan, pembentukan grup tersebut dinamakan Computer-Supported Group Formation (CSGF). Pada umumnya, pembentukan kelompok terdiri dari 3 tahap [14], yaitu :

1. inialisasi proses pembentukan kelompok,
2. pengidentifikasian hubungan dari siswa yang akan dikelompokkan,
3. negosiasi dengan siswa yang akan dibentuk kelompoknya.

Sistem CSGF yang baik adalah sistem yang dapat mendukung semua tahapan diatas, dimana :

1. Tahap inialisasi dilakukan oleh sistem berdasarkan pengetahuan sistem tentang siswa dan keadaan dari siswa tersebut.
2. Identifikasi dari calon pasangan siswa di dalam grup yang akan dibentuk dapat dilakukan dengan menyediakan daftar siswa yang memenuhi persyaratan dari rancangan yang sebelumnya telah dibuat.
3. Negosiasi merupakan proses yang sedikit kompleks, namun dapat dilakukan dengan memberikan kesempatan pada siswa untuk menyetarakan tujuan dan preferensi dari semua individu yang berpartisipasi sekaligus mendefinisikan kepentingan kelompok.

Dalam mencari tahu informasi tentang siswa agar sistem CSGF dapat diinisialisasi, perlu dibuat sebuah parameter yang dapat diproses dan mewakili karakteristik siswa tersebut ketika akan dimasukkan ke dalam grup. Parameter dalam tugas akhir ini adalah average performance siswa.

### 2.2.1 Grup Homogen dan Heterogen

Pengelompokkan dalam CSGF secara garis besar terbagi menjadi 2, yaitu secara homogen yang berdasarkan kesamaan karakteristik siswa yang tidak sering berubah, dan pengelompokkan heterogen yang berdasarkan profil siswa seperti umur, jenis kelamin, etnis, maupun faktor-faktor lain.

Masing – masing jenis pembentukan memiliki kelebihan dan kekurangannya sendiri. Tetapi pada penelitian kali ini, yang menjadi acuan dalam pembentukan grup adalah nilai dan performansi, sehingga jenis grup yang cocok untuk CSGF dengan nilai performansi sebagai acuan adalah grup homogen. Grup homogen juga memberikan hasil yang lebih baik dalam proses belajar dan *assessment* di dalam grup, dimana grup heterogen hanya mengungguli grup

homogen di dalam aspek manajemen dan *research* di dalam grup tersebut [16].

### 2.2.2 Tipe Grup dengan Classifier Berupa Nilai

Terdapat berbagai tipe dan juga pandangan terhadap bagaimana membentuk grup yang cocok dengan pembelajaran kolaboratif [13]. Tiap-tiap tipe grup memiliki kelebihan dan kekurangannya masing-masing sesuai dengan situasi yang dihadapi. Tipe-tipe grup yang dimaksud antara lain :

1. *Pair-Share*, grup beranggotakan dua orang, dimana memiliki kelebihan yaitu lebih nyaman dan sinergis karena hanya beranggotakan dua orang
2. *Jigsaw*, grup yang anggota didalamnya memiliki tugas untuk belajar sekaligus juga mengajar satu sama lain. Kelebihan yang dimiliki adalah selain mendapat pengetahuan siswa juga dapat mengembangkan skill mengajar mereka.
3. *Split-Class*, grup yang hanya membagi kelas menjadi dua grup besar dimana memiliki kelebihan yaitu siswa dapat memperoleh pendapat yang beragam dari sebuah topik yang sedang diperdebatkan.
4. *Random Group of Three*, pembagian dilakukan secara acak dan beranggotakan tiga orang, dimana memiliki kelebihan yaitu siswa dapat memperoleh feedback yang beragam. Tiga orang cukup untuk menambah luas lingkup pembicaraan dari jenis *Pair-Share*, tetapi juga tidak terlalu besar dan memberi resiko terbaikannya beberapa anggota grup.
5. *Ability Group*, pembagian grup dilakukan dengan cara menggabungkan siswa dengan kemampuan yang mirip. Kelebihan dari *ability group* adalah membuat setiap anggota grup merasa nyaman dalam belajar dan berkomunikasi dengan anggota lainnya karena merasa memiliki tingkat pengetahuan yang sama.

### 2.3 Performance Factor Analysis

*Performance Factor Analysis* adalah salah satu metode *tracing* pengetahuan siswa yang merupakan pengembangan dari *Knowledge Tracing dan Learning Factor Analysis* [1]. Algoritma *Knowledge Tracing* memiliki kelemahan yang cukup fatal, dimana KT memiliki *input* parameter yang tidak dapat membaca soal yang memiliki kompetensi lebih dari satu. Padahal, satu buah soal dalam *assessment* tentu memiliki beberapa kompetensi yang harus dikuasai. Algoritma *Learning Factor Analysis* menutupi kelemahan KT dengan mencari nilai logit dari kemampuan belajar siswa menggunakan formula sebagai berikut :

$$m(i, j \in KCs, n) = \alpha_i + \sum_{j \in KCs} (\beta_j + \gamma_j n_{ij})$$

$$p(m) = \frac{1}{1 + e^{-m}}$$

**Gambar 1: Formulasi Standar LFA (1) dan bentuk probabilitiknya (2) [1]**

Algoritma *Performance Factor Analysis* memiliki tujuan yang serupa dengan LFA, namun PFA mengkonfigurasi ulang LFA dan membuatnya sensitif terhadap indikator terkuat dalam pembelajaran siswa, yaitu performansi [1].

Performansi dianggap menjadi sebuah indikator yang kuat dalam pembelajaran dikarenakan oleh dua alasan [15]. Yang pertama, jumlah jawaban benar dari *assessment* yang dikerjakan siswa sangat mengindikasikan bahwa kemampuan siswa saat ini sudah tinggi, sehingga dengan menelusuri jawaban yang benar akan memperkuat kebenaran dari *tracing* yang dilakukan. Kedua, jawaban yang benar akan lebih mengarah ke pembelajaran yang lebih lagi dibanding jawaban salah, dikarenakan tidak efektifnya prosedur *review* yang diberikan setelah kesalahan terjadi. Tetapi, sensitivitas terhadap jawaban yang salah juga memiliki indikator tersendiri dan menjadikan terbentuknya rasio jawaban benar dan jawaban salah, sehingga model yang menggunakan PFA dapat melacak tingkat pengetahuan dan kemampuan belajar siswa dengan efektif. Dengan memasukkan parameter jawaban benar dan jawaban salah tadi, formula LFA menjadi formula PFA dengan bentuk sebagai berikut [1] :

$$m(i, j \in KCs, s, f) = \sum_{j \in KCs} (\beta_j + \gamma_j s_{ij} + \rho_j f_{ij})$$

**Gambar 2: Formulasi PFA**

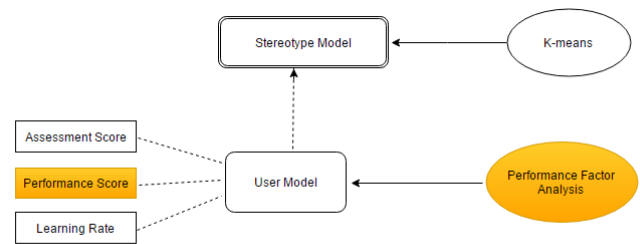
Formulasi PFA hampir serupa dengan LFA, hanya saja PFA menambahkan parameter *s* yaitu parameter yang melacak keberhasilan (jawaban benar) siswa, dan parameter *f* yang melacak kesalahan (jawaban salah) siswa. Variabel  $\gamma$  dan  $\rho$  diatas berfungsi sebagai komponen frekuensi dari hasil *assessment* yang dilakukan, atau dengan kata lain mewakili rasio pengaruh jawaban benar dan jawaban salah terhadap tingkat performansi siswa.

### 3. Analisis Perancangan dan Implementasi

#### 3.1 Deskripsi dan Analisis Sistem

Pada tugas akhir ini, sistem memakai algoritma PFA yang digunakan untuk menentukan nilai performansi siswa. Pembuatan grup lalu dijalankan berdasarkan hasil *Average performance* yang didapat dari algoritma PFA dan disusun mengikuti *stereotype* dari *Average performance*

siswa tersebut. Adapun hubungan antar model di dalam sistem adalah sebagai berikut :



**Gambar 3: Model pada Sistem**  
Sistem memiliki batasan yaitu:

1. Data set merupakan kumpulan nilai praktikum dari modul 1 – 10, nilai UTS, dan nilai UAS mahasiswa yang mengambil mata kuliah Dasar Algoritma dan Pemrograman tahun ajaran 2015 – 2016. Data yang diambil berupa data mahasiswa yang memiliki nilai lengkap dan dimodifikasi sesuai kebutuhan penelitian.
2. Parameter *Performance Factor Analysis* yaitu  $\beta$ ,  $\gamma$ , dan  $\rho$  ditentukan sendiri nilainya sesuai dengan batasan yaitu 0 – 0.5, 0 – 0.5, dan 0 – 0.3 untuk setiap parameter tersebut [1][7].
3. Grup yang dibentuk beranggotakan 3 orang sehingga apabila saat proses pembentukan grup tersisa 2 orang maka kedua siswa tersebut tidak dapat membentuk sebuah grup.
4. Grup hanya berisi siswa yang berada di dalam satu *average performance stereotypes* yang telah dibentuk sebelumnya dan tidak bisa mengambil siswa yang *stereotypes*-nya berbeda untuk melengkapi grup.

#### 3.2 User Model

*User model* pada sistem ini merupakan tempat penyimpanan data berupa nilai – nilai yang didapat dari *dataset* dan juga untuk menyimpan hasil perhitungan performansi dan *average performance* yang didapat dari inferensi *Performance Factor Analysis*.

##### 3.2.1 Atribut User Model

*User model* pada sistem ini berisi nilai spesifik yang dimiliki oleh *user* dari *dataset* yang dipakai. Pada sistem ini nilai – nilai tersebut akan menjadi parameter yang digunakan pada *Performance Factor Analysis* dan beberapa atribut tambahan untuk mendukung sistem nantinya dalam proses pembentukan grup. Atribut – atribut yang terdapat pada *user model* dapat dilihat pada Tabel 1 :

**Tabel 1: Atribut User Model**

Nama Atribut	Keterangan	Nilai dan Tipe Data
<b>Knowledge base</b>	atribut yang merepresentasikan pemahaman awal siswa terhadap materi keseluruhan.	Double. Rentang nilai: 0.0 – 1.0; tidak sama dengan 1. Semakin tinggi maka semakin paham.
<b>Assessment Score</b>	Atribut yang merepresentasikan kesuksesan (skor) siswa pada <i>assessment</i> yang diberikan	Array of Integer. Rentang nilai: 0 – 300. Semakin tinggi maka semakin baik.
<b>Performance Score</b>	Hasil dari inferensi <i>Performance Factor Analysis</i> yang merepresentasikan kemampuan belajar siswa.	Array of Double. Rentang nilai : 0 – 1. Semakin tinggi maka semakin baik.
<b>Average performance</b>	Rata – rata performansi siswa yang didapat dari <i>performance score</i> dari tiap <i>assessment</i> .	Double. Rentang nilai : 0 – 1. Semakin tinggi maka semakin baik.

### 3.3 Performance Factor Analysis

*Performance Factor Analysis* (PFA) merupakan metode inferensi *user model* dimana setiap *assessment* yang dilakukan akan dihitung performansinya untuk menentukan nilai *average performance* siswa tersebut. PFA sepenuhnya mengukur performansi siswa berdasarkan *assessment* yang diberikan tanpa mempertimbangkan performansi di *assessment* sebelumnya dan *prior knowledge* yang dimiliki siswa tersebut.

#### 3.3.1 Parameter Performance Factor Analysis

*Performance Factor Analysis* memiliki 5 parameter utama dimana parameter tersebut akan menentukan nilai performansi dari siswa terhadap sebuah materi. Setiap *assessment* akan memiliki nilai performansinya masing-masing berdasarkan tingkat kesuksesan siswa di *assessment* tersebut. Parameter – parameter yang digunakan pada *Performance Factor Analysis* adalah :

1.  $\beta_i$  atau *item difficulty* adalah tingkat kesulitan dari sebuah *assessment*.
2.  $s_i$  atau *success* adalah nilai kesuksesan dari *assessment* yang dilakukan atau dengan kata lain poin yang didapat dari jawaban benar.
3.  $f_i$  atau *failure* adalah nilai kegagalan dari *assessment* yang dilakukan atau poin yang gagal didapatkan dari *assessment* tersebut.
4.  $Y_i$  adalah variabel frekuensi dari  $s_i$  dimana merepresentasikan bobot *learning* yang didapat dari kesuksesan yang diperoleh.
5.  $\rho_i$  adalah variabel frekuensi dari  $f_i$  dimana merepresentasikan bobot *learning* yang didapat dari kegagalan yang terjadi.

Setiap parameter memiliki cara penentuan yang berbeda – beda. Berikut adalah penentuan nilai parameter – parameter tersebut pada sistem ini :

1.  $\beta_i$ ,  $Y_i$ , dan  $\rho_i$  adalah nilai yang merepresentasikan tingkat kesulitan sebuah *assessment*. Karena penelitian berfokus kepada pembentukan grup dan ketiga parameter tersebut lebih untuk fungsi adaptif, maka di penelitian ini semua *assessment* yang terjadi dianggap memiliki tingkat kesulitan yang sama, sehingga ketiga parameter diatas memiliki nilai konstan untuk setiap *assessment*, dan berada di rentang 0 – 0.5, 0 – 0.5, 0 – 0.3 untuk setiap parameter secara berurutan [1][7].

2.  $s_i$  dan  $f_i$  adalah nilai yang didapat dari nilai *assessment* yang dilakukan dimana rentang nilainya adalah 0 – 1 [1]. Sehingga, nilai  $s_i$  dapat dicari dari jumlah poin yang berhasil didapatkan dibagi dengan jumlah poin maksimal, dan  $f_i$  dapat dicari dari jumlah poin yang gagal didapatkan dibagi dengan jumlah poin maksimal.

#### 3.3.1 Parameter Performance Factor Analysis

*Performance Factor Analysis* digunakan untuk mengukur nilai performansi siswa untuk setiap *assessment* yang diberikan. Setiap nilai yang dimiliki siswa merepresentasikan satu kali *assessment*, dan nilai tersebut menjadi masukan dalam algoritma *Performance Factor Analysis* sebagai poin yang berhasil didapatkan (*success*). Formulasi PFA yang digunakan dalam proses inferensi dapat dilihat di gambar 3.2.

$$m = \beta + (y \cdot (\frac{\text{assessmentScore}}{\text{maxScore}})) + \rho \cdot (\frac{\text{maxScore} - \text{assessmentScore}}{\text{maxScore}})$$

Gambar 3: Formulasi *Performance Factor Analysis*

Dengan menggunakan formula diatas, maka setiap *assessment* yang dilakukan dapat diukur performansinya untuk dapat menghasilkan nilai *average performance*. *Average performance* didapat dengan mencari nilai *mean* dari seluruh nilai performansi yang telah didapat sebelumnya [8].

#### 3.4 Stereotype Model

*Stereotype model* adalah sebuah pengelompokan *user* di dalam *user model* menggunakan sebuah atribut yang dijadikan ciri khas (*stereotype*) [8]. Pada tugas akhir ini, atribut dalam *user model* yang digunakan sebagai *stereotypes* adalah *knowledge base* dan *average performance*. Dari kedua atribut tersebut, kumpulan *user* dapat dikelompokkan menjadi 4 *stereotypes*, yaitu *Low Skill/Low Learning* (LS/LL), *Low Skill/High Learning* (LS/HL), *High Skill/Low Learning* (HS/LL), dan *High Skill/High Learning* (HS/HL) [8].

Seorang siswa masuk ke dalam *stereotype Low Skill* dan *High Skill* ditentukan dari seberapa tinggi *knowledge base* yang dimiliki siswa tersebut, sedangkan *Low Learning* dan *High Learning* ditentukan dari nilai *average*

performance yang didapat dari inferensi *Performance Factor Analysis* sebelumnya.

Untuk menentukan pembagian *range* dari keempat *stereotypes* diatas tidak bisa dilakukan dengan menetapkan *range* yang konstan, karena pada kasus nyata, setiap kelas memiliki standar dan persebaran nilai yang berbeda. Sehingga apabila menetapkan suatu *range* yang konstan, maka akan muncul kemungkinan bahwa seluruh siswa di kelas

tersebut masuk ke dalam *stereotypes* yang sama. Oleh karena itu, untuk menghasilkan pembagian yang merata pengelompokkan *stereotypes* perlu dilakukan menggunakan metode *clustering*, dimana pada tugas akhir ini metode *clustering* yang digunakan adalah *K-means clustering*.

### 3.4.1 Implementasi K-means Clustering terhadap Stereotype Model

*K-means clustering* adalah salah satu metode *clustering* yang membagi data tanpa perlu adanya label kelas. *K-means* dapat diimplementasikan dengan mudah dan waktu yang dibutuhkan untuk menjalankan pembelajaran relatif cepat [15]. Pada tugas akhir ini, spesifikasi *K-means* yang digunakan adalah sebagai berikut :

1. Jumlah *centroid* yang dipakai sesuai dengan jumlah *stereotypes* yang ada, yaitu 4 buah ( $k=4$ ).
2. Pemilihan *centroid* bersifat *random*.
3. Pengukuran jarak titik data terhadap *centroid* menggunakan rumus jarak *Euclidean*.

## 4. Pengujian dan Analisis

### 4.1 Strategi Pengujian

Pengujian sistem dilakukan dengan cara mengubah parameter – parameter *Performance Factor Analysis* dengan rentang nilai sesuai dengan yang telah ditentukan pada perancangan. Pada setiap percobaan dilihat homogenitas grup yang dibentuk dengan cara menjadikan siswa yang memiliki nilai median pada grup sebagai *centroid* sementara dan menghitung jarak 2 siswa lainnya dengan menggunakan *Euclidean distance* untuk mengetahui seberapa dekat kemampuan siswa yang ada di dalam grup tersebut. Semakin kecil nilainya maka semakin mirip kemampuan tiap siswa di dalam grup tersebut, sehingga semakin homogen.

### 4.1 Skenario Pengujian

Dalam pengujian sistem CSGF ini dilakukan dengan menguji parameter *Performance Factor Analysis* yaitu:

1. Nilai  $\beta$
2. Nilai  $\gamma$
3. Nilai  $\rho$
4. Inisialisasi *centroid* pada K-means

Dimana mengikuti inisialisasi *centroid* dan parameter seperti berikut :

Tabel 2: Uji Parameter

$\beta$	$\gamma$	$\rho$
0.5	0.5	0.3
0.4	0.5	0.3
0.3	0.5	0.3
0.2	0.5	0.3
0.1	0.5	0.3

(1) Uji Parameter  $\beta$

$\beta$	$\gamma$	$\rho$
0.5	0.4	0.3
0.5	0.3	0.3
0.5	0.2	0.3
0.5	0.1	0.3

(2) Uji Parameter  $\gamma$

$\beta$	$\gamma$	$\rho$
0.5	0.5	0.2
0.5	0.5	0.15
0.5	0.5	0.1
0.5	0.5	0.05

(3) Uji Parameter  $\rho$

Tabel 3: Inisialisasi Centroid

Centroid	x	y
1	0.764371839	0.159159912
2	0.826460138	0.350562263
3	0.379277662	0.587033992
4	0.646979043	0.325248367

(a) Inisialisasi pertama

Centroid	x	y
1	0.544479947	0.316789951
2	0.392106791	0.174991467
3	0.401783619	0.132537154
4	0.13476611	0.421490435

(b) Inisialisasi kedua

Centroid	x	y
1	0.969326642	0.333808694
2	0.072774189	0.258184287
3	0.059068169	0.439114876
4	0.692688357	0.518028928

(c) Inisialisasi ketiga

Centroid	x	y
1	0.492174562	0.398551338
2	0.771341818	0.58986462
3	0.200890723	0.217654304
4	0.901131104	0.832192852

(d) Inisialisasi keempat

Dalam menentukan nilai homogenitas, digunakan rumus sebagai berikut :

$$homogeneity = \left| \frac{(d2 - d1) + (d3 - d2)}{2} \right|$$

Gambar 4: Formulasi Homogenitas Grup

## 4.2 Hasil Pengujian dan Analisis

Berikut adalah Homogenitas yang dihasilkan menggunakan inisialisasi *centroid* yang pertama dan kedua :

Skenario	Homogenitas Rata-rata
$\beta=0.5, \gamma=0.5, \rho=0.3$	0.033821826
$\beta=0.4, \gamma=0.5, \rho=0.3$	0.037916548
$\beta=0.3, \gamma=0.5, \rho=0.3$	0.043231261
$\beta=0.2, \gamma=0.5, \rho=0.3$	0.050224229
$\beta=0.1, \gamma=0.5, \rho=0.3$	0.069972635

(1) Hasil uji Parameter  $\beta$

Skenario	Homogenitas Rata-rata
$\beta=0.5, \gamma=0.4, \rho=0.3$	0.033233325
$\beta=0.5, \gamma=0.3, \rho=0.3$	0.034355752
$\beta=0.5, \gamma=0.2, \rho=0.3$	0.040677799
$\beta=0.5, \gamma=0.1, \rho=0.3$	0.048517062

(2) Hasil uji Parameter  $\gamma$

Skenario	Homogenitas Rata-rata
$\beta=0.5, \gamma=0.5, \rho=0.2$	0.039173175
$\beta=0.5, \gamma=0.5, \rho=0.15$	0.042030919
$\beta=0.5, \gamma=0.5, \rho=0.10$	0.044935831
$\beta=0.5, \gamma=0.5, \rho=0.05$	0.047968283

(3) Hasil uji Parameter  $\rho$

Gambar 5: Hasil pengujian *centroid* pertama

Skenario	Homogenitas Rata-rata
$\beta=0.5, \gamma=0.5, \rho=0.3$	0.086039302
$\beta=0.4, \gamma=0.5, \rho=0.3$	0.086039302
$\beta=0.3, \gamma=0.5, \rho=0.3$	0.086039302
$\beta=0.2, \gamma=0.5, \rho=0.3$	0.086039302
$\beta=0.1, \gamma=0.5, \rho=0.3$	0.086039302

(1) Hasil uji Parameter  $\beta$

Skenario	Homogenitas Rata-rata
$\beta=0.5, \gamma=0.4, \rho=0.3$	0.086034106
$\beta=0.5, \gamma=0.3, \rho=0.3$	0.086348889
$\beta=0.5, \gamma=0.2, \rho=0.3$	0.086381063
$\beta=0.5, \gamma=0.1, \rho=0.3$	0.086393021

(2) Hasil uji Parameter  $\gamma$

Skenario	Homogenitas Rata-rata
$\beta=0.5, \gamma=0.5, \rho=0.2$	0.085863706
$\beta=0.5, \gamma=0.5, \rho=0.15$	0.085869762
$\beta=0.5, \gamma=0.5, \rho=0.10$	0.085887956
$\beta=0.5, \gamma=0.5, \rho=0.05$	0.085919193

(3) Hasil uji Parameter  $\rho$

Gambar 6: Hasil pengujian *centroid* kedua

Dapat dilihat bahwa parameter *default* menghasilkan nilai rata – rata homogenitas grup yang paling kecil, atau dengan kata lain memiliki hasil grup yang paling homogen. Pada skenario ini juga terlihat bahwa semakin kecil nilai dari parameter yang diuji maka semakin besar nilai homogenitas yang dihasilkan, meskipun tidak menghasilkan perbedaan yang begitu besar.

Dari skenario kedua dapat dilihat bahwa pengaruh parameter yang digunakan terhadap homogenitas masih serupa dengan skenario pertama. Dimana parameter default masih memiliki homogenitas terbaik

. Tetapi, apabila dibandingkan dengan skenario pertama, nilai homogenitas di skenario kedua memiliki perbedaan yang cukup signifikan, dimana skenario pertama menghasilkan nilai homogenitas yang jauh lebih baik dibanding skenario kedua. Perubahan tersebut dihasilkan dari perubahan inisialisasi *centroid* pada K-means.

Adapun grup yang memiliki homogenitas terbaik dari percobaan yang dilakukan adalah sebagai berikut :

Grup	Siswa 1	Siswa 2	Siswa 3	Homogenitas
1	S001	S003	S004	0.060358767
2	S005	S006	S007	0.007602855
3	S008	S009	S010	0.011273367
4	S011	S012	S015	0.003477806
5	S017	S018	S019	0.003692796
6	S023	S024	S029	0.013569735
7	S031	S036	S037	0.009630662
8	S040	S042	S043	0.069459544
9	S052	S053	S056	0.010327312
10	S002	S013	S014	0.078002101
11	S016	S020	S021	0.030265909
12	S022	S025	S026	0.020846246
13	S027	S028	S030	0.008088925
14	S032	S033	S034	0.135190848
15	S035	S038	S039	0.054196587
16	S041	S044	S045	0.034204647
17	S046	S047	S048	0.039916455
18	S049	S050	S051	0.033233325
Orphan	S054	S055	null	null

Gambar 7: Hasil grup yang memiliki homogenitas terbaik

Setelah diketahui bahwa inisialisasi *centroid* memiliki pengaruh yang cukup besar terhadap homogenitas, maka dilakukan kembali percobaan sebanyak 30 kali menggunakan inisialisasi *centroid* yang dihasilkan dari fungsi *random* dan juga menggunakan parameter PFA *default*. Hasil dari percobaan dapat dilihat di tabel berikut :

Iterasi ke-	Homogenitas rata - rata
1	0.062121415
2	0.073737733
3	0.025133113
4	0.087292049
5	0.04812753
6	0.087292049
7	0.070381638
8	0.024733841
9	0.049077519
10	0.086341063
11	0.087292049
12	0.077791298
13	0.045865038
14	0.097578788
15	0.061847866
16	0.097578788
17	0.057783975
18	0.039131709
19	0.086341063
20	0.087292049
21	0.050260582
22	0.019031707
23	0.067756953
24	0.065046541
25	0.087292049
26	0.050642749
27	0.087292049
28	0.065655244
29	0.040694269
30	0.026126219

Gambar 8: Nilai minimum dan maksimum homogenitas yang didapat

Dari iterasi tersebut didapat nilai minimum dan nilai maksimum homogenitas yang dapat dicapai dengan menggunakan parameter *default*, yaitu dengan nilai minimum 0.019031707 dan maksimum 0.097578788. Nilai minimum yang dihasilkan ternyata dapat lebih kecil dari skenario percobaan yang dilakukan sebelumnya. Sehingga terdapat hasil pembentukan grup homogen yang lebih sempurna dibanding grup sebelumnya. Adapun inisialisasi *centroid* yang digunakan ketika mencapai nilai minimum adalah sebagai berikut

Centroid	x	y
1	0.423198255	0.591111851
2	0.776102516	0.54448642
3	0.608905852	0.906367021
4	0.54725452	0.171316605

Gambar 9: Nilai *centroid* pada homogenitas terbaik

Sedangkan grup yang dihasilkan oleh inialisasi centroid diatas adalah sebagai berikut:

Grup	Siswa 1	Siswa 2	Siswa 3	Homogenitas
1	S002	S013	S014	0.044113524
2	S016	S020	S021	0.013102688
3	S022	S027	S028	0.010812103
4	S030	S032	S033	0.020707247
5	S034	S038	S039	0.062156911
6	S041	S044	S045	0.016378966
7	S046	S047	S048	0.020920646
8	S049	S050	S051	0.003913499
9	S003	S008	S042	0.049904188
10	S001	S004	S005	0.011486059
11	S006	S007	S009	0.008630289
12	S010	S011	S012	0.013450853
13	S015	S017	S018	0.005902688
14	S019	S023	S024	0.016721139
15	S025	S026	S029	0.023996308
16	S031	S035	S036	0.008710529
17	S037	S040	S043	0.004887706
18	S052	S053	S055	0.006775394
Orphan	S054	S056	null	null

Gambar 9: Hasil grup dengan homogenitas terbaik

#### 4.2 Stereotype dan Orphan Student

Karena K-means tidak memiliki label kelas, maka untuk memberikan *stereotype* pada kelas perlu dilakukan pengamatan.

Pada skenario pertama untuk parameter yang menghasilkan homogenitas terbaik pun menghasilkan cluster yang serupa seperti yang dapat dilihat di gambar berikut :

Skenario	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Iterasi	Jumlah Grup terbentuk	Homogenitas Rata - rata
$\beta=0.5, \gamma=0.4, \rho=0.3$	0	27	29	0	10	18	0.033233325

Gambar 10: Hasil clustering dari grup pertama

sehingga berikut adalah *stereotype* yang dibangun berdasarkan letak *centroid*-nya :

Centroid	x	y	Stereotype 1	Stereotype 2
1	0.646979043	0.159159912	H	L
2	0.764371839	0.870722022	H	H
3	0.385862069	0.86617637	L	H
4	0.631481481	0.325248367	L	L

Gambar 11: Stereotype pada grup pertama

Dari kedua gambar diatas dapat disimpulkan bahwa 27 siswa termasuk kedalam stereotype High Skill/High Learning dan 29 siswa termasuk ke dalam stereotype Low Skill/High Learning.

Pada pengujian kasus siswa yang tidak memiliki kelompok atau orphan student, dengan mengesampingkan 2 orang yang tidak dapat membentuk grup karena jumlah siswa tidak habis dibagi 3, hanya terjadi 1 kali terbentuknya orphan student lebih dari 2 orang, yaitu dengan jumlah *orphan student* 5 seperti pada gambar berikut :

$\beta=0.5, \gamma=0.5, \rho=0.3$	2
$\beta=0.4, \gamma=0.5, \rho=0.3$	5
$\beta=0.3, \gamma=0.5, \rho=0.3$	2
$\beta=0.2, \gamma=0.5, \rho=0.3$	2
$\beta=0.1, \gamma=0.5, \rho=0.3$	2
$\beta=0.5, \gamma=0.4, \rho=0.3$	2
$\beta=0.5, \gamma=0.3, \rho=0.3$	2
$\beta=0.5, \gamma=0.2, \rho=0.3$	2
$\beta=0.5, \gamma=0.1, \rho=0.3$	2
$\beta=0.5, \gamma=0.5, \rho=0.2$	2
$\beta=0.5, \gamma=0.5, \rho=0.15$	2
$\beta=0.5, \gamma=0.5, \rho=0.10$	2
$\beta=0.5, \gamma=0.5, \rho=0.05$	2

## 5. Kesimpulan dan Saran

### 5.1 Kesimpulan

Berdasarkan hasil perancangan, analisis dan pengujian dapat diambil kesimpulan dari penelitian ini sebagai berikut:

1. Sistem CSGF dapat diterapkan dengan menggunakan metode *Performance Factor Analysis* dimana algoritma PFA berfungsi sebagai metode untuk mengukur *performance* siswa dalam menjalani *assessment*. *Stereotype model* dicari dengan menggunakan metode *K-means clustering* dan digunakan sebagai acuan untuk membentuk grup yang homogen dan sesuai dengan tujuan.
2. Homogenitas dari grup yang dibentuk sangat dipengaruhi oleh inialisasi *centroid* yang dilakukan. Semakin baik inialisasi maka hasil akhir pembentukan akan semakin baik juga. Parameter – parameter dalam PFA juga mempengaruhi nilai homogenitas, dimana semakin kecil nilai parameter maka semakin besar nilai homogenitasnya, namun pengaruhnya tidak sebesar inialisasi *centroid*.
3. Sistem CSGF dapat menghasilkan *stereotype model* dengan baik dan cukup konsisten, permasalahan *orphan student* juga sangat langka terjadi sehingga dapat dikatakan bahwa sistem menangani pembentukan grup dengan cukup baik.

### 5.2 Saran

Dalam menentukan nilai parameter – parameter yang digunakan pada sistem masih menggunakan cara independen. Dengan melakukan kajian lebih dalam lagi terhadap penentuan nilai parameter – parameter *Performance Factor Analysis* yaitu  $\beta$ ,  $\gamma$ , dan  $\rho$ , sistem yang dihasilkan akan lebih baik dan lebih akurat. Penentuan *stereotype model* dapat dilakukan dengan metode *clustering* selain *K-means* untuk menghasilkan *model* yang lebih merata dan lebih akurat dibanding *K-means*.

### Daftar Pustaka

- [1] Pavlik, Phillip.L, Hao Cen, and Kenneth R. "Performance Factors Analysis – A New Alternative to Knowledge Tracing." *Human Computer Interaction Institute, Carnegie Mellon University*. USA, 2009.
- [2] Hoppe, H. Ulrich. "The Use of Multiple Student Modeling to Parameterize Group Learning." *University of Duisburg (FRG) D-47048*. Duisburg, 1995.
- [3] Wertsch, James V. "From social interaction to higher psychological processes. A clarification and application of Vygotsky's theory." *Human development* 22.1 (1979) : 1-22
- [4] Plech, Chris, Jonathan Spencer, Jonathan Huang. "Deep Knowledge Tracing." *Stanford University. Khan Academy. Google* (2015) :1-3
- [5] Strijbos, Jan-Willem. "Assessment of (Computer-Supported) Collaborative Learning". *IEEE Transactions On Learning Technologies*, Vol. 4, No. 1. 2011



- [6] Cen, Hao. Kenneth Koedinger, and Brian Junker. *“Learning Factor Analysis – A General Method for Cognitive Model Evaluation and Improvement”*. Carnegie Mellon University, 5000 Forbes, Pittsburgh, PA. USA. 2005
- [7] Gong, Y, Beck, J. E., Heffernan, N. T.: *How to Construct More Accurate Student Models: Comparing and Optimizing Knowledge Tracing and Performance Factor Analysis*. International Journal of Artificial Intelligence in Education. Accepted, 2010.
- [8] Rafferty, Anna N. and Michael Yudelson. *“Applying Learning Factors Analysis to Build Stereotypic Student Models”*. Stanford University, University of Pittsburgh, PA. USA. 2004
- [9] Mitrovic, A., Koedinger, K. R., & Martin, B. *A comparative analysis of cognitive tutoring and constraint-based modeling*. In P. Brusilovsky, A. Corbett, & F. de Rosis (Eds.), *Proceedings of the Ninth International Conference on User Modeling*, 313-322. Berlin: Springer-Verlag. 2003
- [10] Gabelnick, F. and J. MacGregor, R. Matthews, and B.L. Smith. *Learning Communities: Creating Connections Among Students, Faculty and Disciplines*. San Francisco: Jossey Bass New Directions for Teaching and Learning, Number 41, Spring 1990.
- [11] Golub, J. (Ed). *Focus on Collaborative Learning*. Urbana, IL: National Council of Teachers of English, 1988.
- [12] Newell, A., Rosenbloom, P.: *Mechanisms of Skill Acquisition and the Law of Practice*. In Anderson J. (ed.): *Cognitive Skills and Their Acquisition*, Erlbaum Hillsdale NJ (1981)
- [13] Cohen, G. *Designing Groupwork Strategies For The Heterogenous Clasroom*. New York: Teachers College Press, 1986.
- [14] Wessner, M., & Pfister, H. R. “Group formation in computer-supported collaborative learning.” *Proceedings of the 2001 International ACM SIGGROUP Conference on Supporting Group Work*, New York (2001) : ACM, 24–31.
- [15] Russell, S. and P. Norvig, *Artificial Intelligence A Modern Approach*. Upper Saddle River, New Jersey 07458: Pearson Education, Inc., 3 ed., 2010.
- [16] Watson. Larry Michaelsen. Cultural Diversity’s Impact on Group Process and Performance : Comparing Culturally Homogenous and Culturally Diverse Task Group The Academy of Management Journal 36(3):590-602 · June 1993.