

STUDI SIMULASI MENGGUNAKAN *FUZZY C-MEANS* DALAM MENGLASIFIKASI KONSTRUK TES

Rukli

STMIK Lamappapoleonro Soppeng
Ksatria No 56 Soppeng
rukli123@yahoo.com

Abstrak

Tulisan ini memperkenalkan metode *fuzzy c-means* dalam mengklasifikasi konstruk tes. Untuk memverifikasi sifat unidimensional suatu tes biasanya menggunakan analisis faktor sebagai bagian dari statistik parametrik dengan beberapa persyaratan yang ketat sedangkan metode *fuzzy c-means* termasuk metode heuristik yang tidak memerlukan persyaratan yang ketat. Studi simulasi penelitian ini menggunakan dua metode yakni analisis faktor menggunakan program SPSS dan *fuzzy c-means* menggunakan program Matlab. Data simulasi menggunakan tipe data dikotomi dan politomi yang dibangkitkan lewat program Microsoft Office Excel dengan desain 2 kategori, yakni: tiga butir soal dengan banyak peserta tes 10, dan 30 butir soal dengan banyak peserta tes 100. Hasil simulasi menunjukkan bahwa metode *fuzzy c-means* lebih memberikan gambaran pengelompokan secara deskriptif dan dinamis pada semua desain yang telah dibuat dalam memverifikasi unidimensional pada suatu tes.

Kata kunci: *fuzzy c-means*, analisis faktor, unidimensional

SIMULATION STUDY USING FUZZY C-MEANS FOR CLASIFYING TEST CONSTRUCTION

Rukli

STMIK Lamappapoleonro Soppeng
Ksatria No 56 Soppeng
rukli123@yahoo.com

Abstract

This paper introduces the fuzzy c-means method for classifying the test constructs. To verify the unidimensional a test typically uses factor analysis as part of parametric statistics with some strict requirements, while fuzzy c-means methods including method heuristic that do not require strict requirements. Simulation comparison between the method of factor analysis using SPSS program and fuzzy c-means methods using Matlab. Simulation data using data type dichotomy and politomus generated through Microsoft Office Excel programs each with a number of 3 items using the number of participants 10 tests, while the number of 30 test items using the number as many as 100 participants. The simulation results show that the fuzzy c-means method provides a more descriptive and dynamic grouping of all the designs that have been made in verifying the unidimensional as a test.

Keywords: *fuzzy c-means, factor analysis, unidimensional*

Pendahuluan

'Bu, berapa nilai matematika Habibi? Sembilan, Pak Budi'. 'Wah, bagus Bu, anaknya pasti pintar'. Kalimat tersebut kadangkala diucapkan dan didengarkan pada saat pembagian rapor di sekolah. Lalu, apa artinya nilai sembilan tersebut? Pertanyaan tersebut mengarah pada unidimensi tes.

Kata unidimensi suatu tes sering kali 'diperlunak' menjadi kata dominan yang menunjukkan konstruk tes dalam mengukur kemampuan peserta tes. Menurut Hambleton *et al.* (1991: 9), karakteristik butir soal-butir soal yang membentuk tes pada pengukuran psikologi tidak tepat secara eksak berdimensi satu (unidimensi) tetapi hanya dominan pada suatu unjuk kerja. Bila tes mengukur lebih dari satu dimensi atau tidak dominan maka jawaban peserta tes terhadap butir soal akan merupakan kombinasi dari berbagai kemampuan, sehingga sulit untuk mengetahui kontribusi dari setiap butir soal atau dimensi terhadap kemampuan peserta tes. Sejatinya, unidimensi merupakan asumsi yang perlu dipenuhi suatu tes sebelum diaplikasikan dalam mengukur kemampuan peserta tes.

Tes memuat satu, dua atau lebih butir soal yang dimaksudkan untuk mengukur *trait* peserta tes, dimana *trait* sendiri adalah kemampuan laten yang perlu diukur. Butir soal tersebut diturunkan dari indikator sedangkan indikator diturunkan dari konsep dasar yang dioperasionalkan sesuai dengan materi pembelajaran. Tes yang dibangun tadi, misalnya tes matematika dimaksudkan untuk mengukur sejauh mana peserta tes dapat menguasai materi yang telah diajarkan selama beberapa waktu (misalnya, satu semester). Dalam kurun waktu tersebut peserta tes belajar berbagai pokok bahasan atau beberapa sub pokok bahasan misalnya pecahan, penjumlahan, pembagian, pengurangan, dan perkalian.

Pengklasifikasian tersebut mudah dilakukan karena butir soal yang membentuk tes dianggap mengukur satu tingkat penguasaan tertentu, yang dalam tulisan ini mengarah pada tes yang dominan, yakni dominan mengukur kemampuan matematika. Walaupun demikian, bila penjumlahan, pengurangan, perkalian, atau pembagian tersebut dianggap domain yang berbeda maka analisis dan penafsiran yang dilakukan akan berbeda. Jika guru membuat tes prestasi belajar matematika diharapkan mencakup semua

materi tersebut, serta melaporkan hasilnya. Misalnya, Habibi menjawab 8 butir soal dari 10 butir soal pada materi penjumlahan, berarti tingkat penguasaannya 80% termasuk kategori baik pada materi penjumlahan. Jika menjawab 3 butir soal dari 10 butir soal pada materi perkalian berarti tingkat penguasaan 30%, termasuk kategori kurang. Tentu dengan hasil tersebut akan lebih menunjukkan kemampuan spesifik dalam mendiagnosis tingkat kesulitan siswa pada materi tersebut jika dibandingkan dengan nilai 9 yang diperoleh Habibi pada mata pelajaran matematika.

Walaupun tes prestasi belajar lebih menekankan validitas isi daripada validitas konstruk namun bukan berarti validitas konstruk tidak penting terutama pada tes yang cakupan materi yang agak luas. Misalnya, tes kenaikan kelas atau tes akhir dimana dimensi atau faktor dalam tes perlu diketahui sedemikian sehingga penskoran akan lebih akurat dalam mencerminkan keterwakilan materi pelajaran dalam mengukur kemampuan yang lebih spesifik.

Berdasarkan gambaran tersebut, ada beberapa hal yang dapat dipaparkan, misalnya kenapa tingkat penguasaan siswa pada materi pecahan berada kategori baik namun pada materi perkalian berada pada kategori rendah, padahal untuk memahami secara baik pembagian diperlukan penguasaan perkalian? Pertanyaannya, apakah Habibi mendapat nilai 9 pada mata pelajaran matematika dapat mengungkap fakta tersebut? Pertanyaan tersebut sukar untuk dijawab bila menganggap tes matematika itu satu dimensi namun melibatkan berbagai pokok bahasan yang berbeda. Jadi, tes yang dirancang mengukur kemampuan yang cakupannya luas memerlukan kajian lebih lanjut mengenai karakteristik dimensi sebelum melakukan penskoran dengan menggunakan metode tertentu.

Ada beberapa metode yang dapat digunakan dalam menentukan tes yang berdimensi satu. Metode penyelidikan unidimensi dapat dilakukan melalui analisis faktor eksploratori lalu dilanjutkan dengan analisis faktor konfirmatori jika teori yang membangun faktor tersebut 'belum mapan'. Sebaliknya jika teori sudah mapan yang membangun pengklasifikasian faktor tersebut sehingga dapat diturunkan dalam bentuk hipotesis maka dapat menggunakan analisis faktor konfirmatori tanpa melalui analisis faktor eksploratori. Hasil penelitian Wu (2006) menunjukkan model teori

respon butir multidimensional dapat mengekstrak informasi pada butir soal-butir soal tes bentuk *linked* yang tidak dapat dikaji menggunakan analisis faktor. Sedangkan penelitian Yu, *et al.* (2007) dalam membandingkan tiga metode dalam mengevaluasi unidimensional menemukan bahwa metode *tetrad* sukses mengkonfirmasi satu dimensi pada sehim-punan data yang dikonstruksi tunggal dan dua dimensi dari sehim-punan data yang dikombinasikan, walaupun satu butir soal untuk tiap klaster tidak dijelaskan lebih lanjut. Walaupun ketiga metode sama-sama mensyaratkan penilaian yang baik dalam mengevaluasi unidimensional. Dunn-Rankin, *et al.* (2004) menyebutkan ada empat analisis klaster yang sederhana dan tidak memerlukan asumsi, yakni *graphic similarity analysis*, *single and complete linkage klaster ing*, *divisive klaster ing* dan *k-means iterative klaster ing*. Keempat metode tersebut bermanfaat dalam menggambarkan dimensional instrumen khususnya dalam mengidentifikasi butir soal-butir soal yang *do not belong*. Lelli (2002) melakukan perbandingan analisis faktor dengan teori himpunan *fuzzy* menemukan bahwa metode dengan pendekatan *fuzzy* lebih fleksibel dan menjanjikan untuk digunakan dimasa yang akan datang. Oleh karena itu, tujuan penelitian ini adalah mengkaji secara simulasi konstruk suatu tes dengan menggunakan metode *fuzzy c-means* dan metode analisis faktor.

Analisis faktor merupakan analisis komponen utama yang sering digunakan peneliti dalam mengidentifikasi faktor, yakni banyak faktor yang relatif kecil yang dapat digunakan untuk menjelaskan sejumlah besar variabel yang saling berhubungan. Pengharapannya, indikator dalam satu faktor mempunyai korelasi yang tinggi, sedangkan korelasi dengan indikator lain relatif rendah atau kalau bisa tidak ada korelasi sama sekali. Tiap-tiap kelompok dari indikator atau variabel mewakili suatu konstruksi dasar yang disebut faktor.

Matriks loading ditransformasikan agar daya interpretasi faktor lebih nampak kecenderungan pengumpulan atau pemisahan antar indikator. Transformasi dapat dilakukan dengan merotasi matriks tersebut dengan metode tertentu misalnya *varimax*, *quartimax*, *equamax*, *quartimin*, *biquartimin*, *covarimin* atau *oblimin*. Hasil rotasi tersebut akan memetakan setiap indikator asal mempunyai korelasi tinggi dengan faktor tertentu saja sedangkan

faktor yang lain berkorelasi relatif rendah sehingga setiap faktor akan lebih mudah untuk diinterpretasikan.

Analisis faktor terbagi dua, yakni *Exploratory Factor Analysis (EFA)* dan *Confirmatory Factor Analysis (CFA)*. Beberapa peneliti skeptis walaupun masih diperdebatkan pada aturan konvensional untuk mengekstrak faktor dengan menggunakan kriteria Kaiser (Chong Ho Yu, *et al.* 2007; Ledesma, *et al.* 2007). Demikian halnya, pengaplikasian *EFA* dalam memverifikasi dugaan, apakah hipotesis tersebut diterima atau ditolak pada beberapa penelitian menggunakan kriteria beragam.

Program Lisrel menggunakan banyak aturan dalam menetapkan fit atau tidaknya suatu model, dan amat melimpah sehingga beberapa hasil penelitian mengambil keputusan yang berbeda-beda dan sering membingungkan. Susanto (2009) menggunakan nilai *Goodness of Fit Indices (GFI)* yakni suatu ukuran keputusan dalam ketepatan model berdasarkan matriks kovarians. Nilai antara *first order* dan *second order* secara teori tidak konsisten dimana nilai *second order* pada semua masa ujian lebih tinggi jika dibandingkan dengan *first order* bahkan mendekati 1. Selanjutnya *Comparative Fit Index (CFI)* lebih rendah maka dipilih *Root Mean Square Error of Approximation (RMSEA)* dan *GFI*. Pemilihan tersebut tidak disertai dengan argumen dan gambaran yang menyakinkan sehingga yang nampak suatu patokan memilih yang cocok dan menghindari dari yang tidak cocok. Susongko (2009) melakukan uji unidimensi dengan melibatkan beberapa indikator, misalnya perbandingan antara nilai Chi Kuadrat dengan derajat bebas hanya berpedoman pada Wheaton (1977) sebesar 5, namun seandainya menggunakan aturan Carmins dan Molver (1981) sama dengan 2 (Ghozali & Fuad, 2008: 327) maka kesimpulan yang diambil akan lain yakni modelnya tidak fit. Selanjutnya penelitian Sunkoco (2009) pada uji model diantara indikator uji kecocokan model yang ada hanya satu nilai yang kritis, yakni nilai *Adjusted Goodness of Fit Indices (AGFI)* sama dengan 0,90 yang persis sama dengan nilai batasan tersebut namun dikatakan seharusnya lebih besar dari 0,90 namun tetap dinyatakan memenuhi model. Berdasarkan kajian di atas nampak begitu ragamnya cara penafsiran dan proses dalam mencapai tingkat kecocokan dengan model namun kadang melupakan dasar argumentasi.

Bollen dan Long (1993), MacCallum (1990), Mulaik, *et al.* (1989) dan Steiger (1991) dalam Gozali dan Fuad (2008) mengatakan bahwa evaluasi model *fit* pada SEM merupakan masalah yang belum terpecahkan dan sangat sulit. Lebih lanjut, Jackson, *et al.* (2005) mengemukakan bahwa pilihan antara ukuran sampel, banyak parameter, dan derajat kebebasan sebagai patokan kecocokan model masih dicari yang pada akhirnya jika tidak cocok maka dapat memilih beberapa alternatif, sebagai berikut: Pertama, meninggalkan keseluruhan projek dimana tidak ada generalisasi direkomendasikan. Kedua, model cadangan yang lain sebagai pengganti berdasarkan teori dan pengetahuannya jadi acuan. Ketiga memodifikasi model asli supaya lebih cocok dengan berbagai konsekuensi-konsekuensi. Dari analisis yang ada ternyata alternatif ketiga lebih menonjol ketimbang alternatif yang lain, walaupun pemilihan tersebut kadang kala memodifikasi tanpa dasar teori yang kuat.

Keputusan tersebut dapat dipahami namun kurang objektif karena ada kecenderungan pemilihan kriteria tersebut lebih kepada sesuai dan tidak sesuai namun analisis logikanya tidak nampak. Akan muncul dilematis dimana ada dua kriteria yang cocok namun yang lain tidak cocok atau variasi lainnya sehingga ada unsur *like and dislike* sehingga yang cocok saja yang diambil tanpa memberikan penjelasan detail yang tidak cocok. Keputusan tersebut akan menyebabkan proses analisis selanjutnya kabur yang pada akhirnya dipertanyakan proses logikanya. Salah satu metode sederhana dan mudah diterapkan dalam menggambarkan unidimensional tes adalah metode *fuzzy c-means* (FCM) berdasarkan aturan logika *fuzzy*.

FCM merupakan metode penetapan pengelompokan secara optimal dalam satu ruang vektor. Tiap data ditentukan oleh derajat keanggotaan yang berbeda antara 0 dan 1 yakni jarak antar vektor pada ruang normal Euclidean. Tingkat keberadaan data dalam suatu kelas atau klaster ditentukan oleh derajat keanggotaannya. Vektor-vektor dalam satu klaster mempunyai derajat kesamaan yang tinggi demikian juga untuk klaster yang lain, sebaliknya antar klaster derajat kesamaan anggotanya rendah atau tidak ada.

Aturan logikanya, tentukan pusat klaster yang akan menandai lokasi rata-rata untuk tiap-tiap klaster. Kondisi awal sebagai pusat klaster ini

masih belum akurat. Tiap-tiap data memiliki derajat keanggotaan pada tiap-tiap kluster. Dengan cara memperbaiki pusat kluster dan nilai keanggotaan tiap-tiap data secara berulang, maka dapat dilihat kecenderungan pusat kluster akan menuju lokasi yang tepat. Perulangan ini didasarkan pada minimasi fungsi obyektif.

Zimmerman (1993) dan Sen-Chi Yu (2008) mengemukakan prosedur perhitungan FCM secara lengkap, yakni: (1) tentukan matriks X berukuran $n \times m$, dengan n = banyak data yang akan di kluster; dan m = banyak variabel (kriteria). Penentuan n dan m disesuaikan dengan kondisi data di lapangan. (2) tentukan banyak kluster yang akan dibentuk dimana banyak kluster lebih besar atau sama dengan 2 ($k \geq 2$). (3) tentukanlah pangkatnya dimana bernilai lebih besar 1 (pembobot $w > 1$). (4) tentukan maksimum iterasi yang diinginkan. (5) tentukan kriteria penghentian (ϵ = nilai positif yang sangat kecil). (6) bentuklah matriks partisi awal U (derajat keanggotaan dalam kluster) dengan matriks partisi awal biasanya dibuat secara acak.

$$\begin{bmatrix} \mu_{11}(x_1) & \mu_{12}(x_2) & \dots & \mu_{1n}(x_n) \\ \mu_{21}(x_1) & \mu_{22}(x_2) & \dots & \mu_{2n}(x_n) \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & & \cdot \\ \mu_{c1}(x_1) & \mu_{c2}(x_2) & \dots & \mu_{cn}(x_n) \end{bmatrix}$$

(7) hitung pusat *klaster* V untuk setiap *klaster*

$$v_{ij} = \frac{\sum_{k=1}^n (\mu_{ik})^w \cdot x_{kj}}{\sum_{k=1}^n (\mu_{ik})^w}$$

(8) memperbaiki derajat keanggotaan setiap data pada setiap kluster (perbaiki matriks partisi) dengan menggunakan persamaan:

$$\mu_{ik} = \left(\sum_{j=1}^c \left[\frac{d_{ik}}{d_{jk}} \right]^{2/(w-1)} \right)^{-1} \text{ dimana}$$

$$d_{ik} = d(x_k - v_i) = \left(\sum_{j=1}^m (x_{kj} - v_{ij}) \right)^{1/2}$$

(9) tentukanlah kriteria penghentian iterasi, yaitu perubahan matriks partisi pada iterasi sekarang dan iterasi sebelumnya sebesar delta: $\Delta \equiv \|U^t - U^{t-1}\|$. Apabila $\Delta < \epsilon$ maka iterasi dihentikan jika tidak lanjutkan menghitung kembali pusat kluster ke-k.

Pada kondisi data riil yang berukuran kecil aturan di atas dapat dilakukan secara manual dengan perhitungan yang cukup kompleks sebagaimana yang dilakukan oleh Teknomo (2007), namun untuk data berukuran besar dapat menggunakan program tertentu, misalnya program Matlab. Aturan umum program Matlab perlu diperhatikan dalam membuat kluster menyangkut *linkage create hierarchical kluster tree*. $Z = \text{linkage}(y, \text{method})$ adalah pembuatan sebuah kluster secara hirarki dengan menggunakan algoritma tunggal linkage (*the single linkage algorithm*) atau lainnya disesuaikan dengan data. *The input y* adalah matriks jarak yang dibangkitkan dengan menggunakan PDIST. Pemilihan metode dapat menggunakan *single* untuk *nearest distance*, *complete* untuk *furthest distance*, *average* untuk *unweighted average distance (upgma)* atau rata-rata kelompok, dan *kluster tree* harus monotonik, “*weighted*” untuk *weighted average distance (wpgma)*, *centroid* untuk *unweighted center of mass distance (upgmc)* dan *kluster tree* harus monotonik, *median* untuk *weighted center of mass distance (wpgmc)* dan *kluster tree* harus monotonik, dan *ward* untuk *inner squared distance (min variance algorithm)*. Pemilihan tersebut bergantung pada objek yang diteliti, namun yang paling sering digunakan adalah *single* yaitu menggunakan jarak terdekat.

Metode Penelitian

Penelitian ini menggunakan metode simulasi. Dilakukan dengan menggunakan data dikotomi (1,0) dan politomi (1,2,3,4,5) dengan desain 2 kategori, yakni panjang tes 3 butir untuk 30 peserta tes panjang tes 10 butir untuk 100 peserta tes. Metode analisis faktor yang digunakan adalah *CFA* lewat program SPSS sedangkan analisis *fuzzy c means* menggunakan program Matlab dimana metode *linkage* yang digunakan adalah *single*, *award*, *complete*, dan *average*. Data simulasi berupa data dikotomi dibangkitkan menggunakan program Microsoft Excel dengan perintah $=\text{round}(\text{rand}()*1+0;0)$. Simulasi dilakukan dua kali dengan ukuran sampel yang berbeda, yakni simulasi I (3 butir dengan 30 peserta tes) dan simulasi II (10 butir dengan 100 peserta tes). Tiap simulasi tersebut dianalisis menggunakan kedua metode *CFA* dan *fuzzy c means*.

Demikian pula pada data politomi, data politomi dibangkitkan menggunakan program Excel dengan perintah $=\text{round}(\text{rand}()*4+1;0)$. Simulasi dilakukan dua kali dengan ukuran sampel yang berbeda, yakni simulasi I yang terdiri atas 3 butir soal dengan 30 peserta tes dan simulasi II yang terdiri atas 10 butir soal dengan 100 peserta tes. Data tiap simulasi tersebut dianalisis menggunakan kedua metode tersebut sebagai berikut:

Hasil Penelitian dan Pembahasan

Simulasi I: 3 Butir dan 30 Peserta Tes menggunakan *CFA* dan *Fuzzy C Means*

Hasil analisis deskriptif ditunjukkan pada Tabel 1, tingkat kecukupan ukuran sampel sebesar 42% (termasuk rendah) karena banyak peserta tes hanya 30. Kecocokan data dengan model menggunakan Tes Batlley lewat pendekatan Chi Kuadrat terlihat data berdistribusi normal dimana nilai sig = 0,618 lebih besar daripada 0,05.

Tabel 1. KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy		,428
Bartlett's Test of Sphericity	Approx. Chi-Square	1,786
	Df	3
	Sig.	,618

Tabel 2 menunjukkan butir soal 1 dan butir soal 3 berada pada komponen 1 sedangkan butir soal kedua berada pada komponen 2. Gambar 1 menunjukkan plot tiga butir soal.

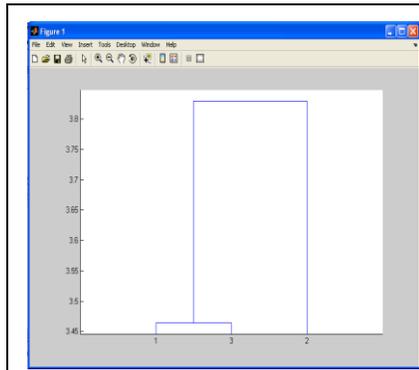
Tabel 2. Matriks Komponen yang Telah Dirotasi

Rotated Component Matrix ^a		
	Component	
	1	2
b1	,811	-,317
b2	-,088	,879
b3	,699	,500

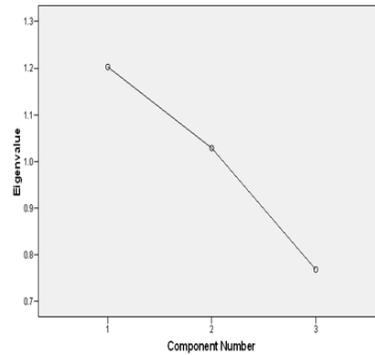
Extraction Method: Principal Component Analysis.
 Rotation Method: Varimax with Kaiser Normalization.
 a. Rotation converged in 3 iterations.

Analisis *Fuzzy* C-Means dengan menggunakan program Matlab yakni `>> X = xlsread('D:\butir soal3_30'); >> x= X'` (transpose matriks dari baris jadi kolom dan kolom jadi baris, karena matlab menganggap baris itu adalah butir soal sedangkan spss menganggap kolom adalah butir soal); `>> y=pdist(x); >> Z=linkage(y,'ward'); >> H=dendrogram(Z)`.

Gambar 2 menunjukkan butir soal 1 dan butir soal 3 berada dalam satu komponen dengan menggunakan metode *linkage ward*, sedangkan butir soal 2 berada pada komponen lain. Kecenderungan pola tersebut terjadi juga pada metode *linkage* yang lain.



Gambar 1. Plot Tiga Butir soal



Gambar 2. Dendrogram Tiga Butir soal

Berdasarkan analisis tersebut dengan melibatkan 3 butir soal dengan ukuran sampel 30, kedua metode mempunyai pola yang sama dalam membuat kategori komponen. Walaupun informasi yang diberikan pada analisis faktor lebih banyak, misalnya nilai kecukupan ukuran sampel dan normalisasi serta variansi namun hal tersebut tidak terlalu signifikan digunakan terutama dalam kajian eksploratori.

Simulasi II: 10 Butir Soal dan 100 Peserta Tes menggunakan *CFA* dan *Fuzzy C Means*

Tabel 3 menunjukkan tingkat kecukupan ukuran sampel sebesar 46% (termasuk rendah) dengan banyak peserta tes sebesar 100. Peningkatan ukuran sampel 10 menjadi 100 menghasilkan selisih tingkat kecukupan ukuran sampel hanya 2%, sedangkan kecocokan data dengan model dengan menggunakan Tes Bartlett lewat pendekatan Chi Kuadrat terlihat data berdistribusi normal dimana nilai sig = 0,369 lebih besar dari 0,05. Kenaikan ukuran sampel dari 10 ke 100 menurunkan nilai sig = 0,249. Penurunan tersebut memberi informasi bahwa tingkat perbedaan 100 data yang dibangkitkan mempunyai variansi yang lebih besar jika dibandingkan 10. Walaupun kecenderungan tersebut tidak dapat digeneralisasikan karena

secara teori peningkatan ukuran sampel yang mendekati ukuran populasi selayaknya cenderung berdistribusi normal.

Tabel 3. KMO and Bartlett's Test

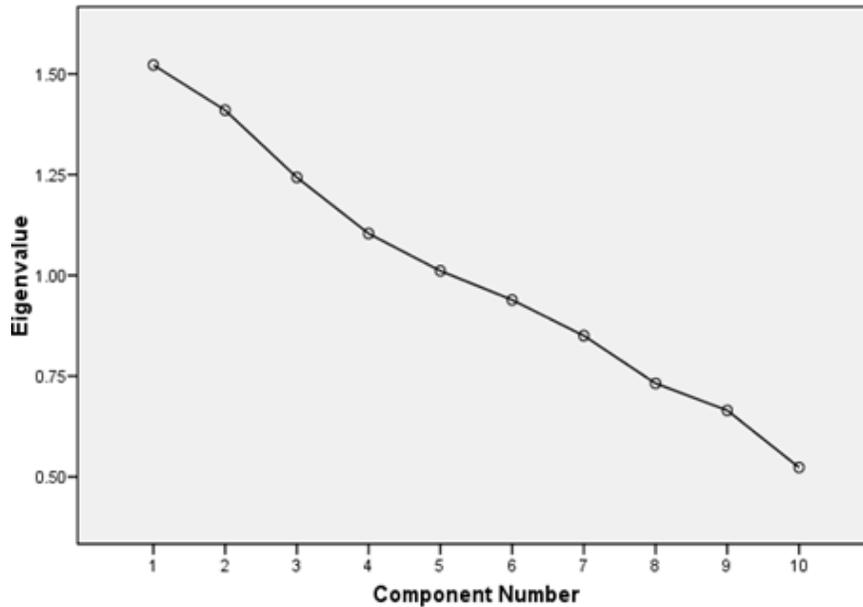
Kaiser-Meyer-Olkin Measure of Sampling Adequacy		,464
Bartlett's Test of Sphericity	Approx. Chi-Square	47,545
	Df	45
	Sig.	,369

Tabel 4 menunjukkan 10 butir soal mengelompok ke dalam lima komponen, komponen pertama memuat butir soal 3, 8, dan 10, komponen dua memuat butir soal 1, komponen tiga memuat butir soal 2 dan 7, komponen empat memuat butir soal 6, dan komponen lima memuat butir soal 8 sedangkan butir soal 4 dan 5 dan 9 mempunyai nilai negatif masing-masing berada pada komponen empat, tiga, dan lima. Butir soal 8 mengukur dua komponen yakni komponen satu dan lima.

Tabel 4. Hasil Matriks Komponen Setelah Dilakukan Rotasi

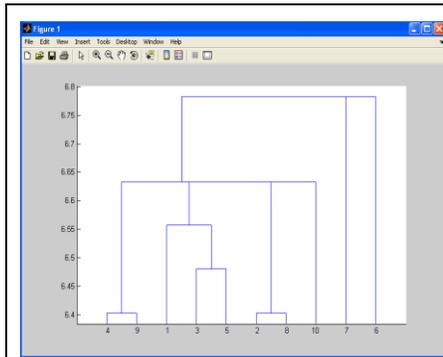
	Component				
	1	2	3	4	5
b1	-,051	,731	,094	,049	,054
b2	-,190	-,526	,350	-,038	,271
b3	,680	,081	-,051	-,162	-,145
b4	-,048	,054	-,042	-,855	,123
b5	,208	-,306	-,726	,214	,149
b6	-,346	,331	-,099	,534	,114
b7	,137	-,200	,743	,154	,113
b8	,522	,351	-,068	-,052	,445
b9	,088	,064	-,033	,051	-,885
b10	,629	-,402	,055	,279	,044

Gambar 3 di bawah menunjukkan pengelompokan tabel 4 dimana 10 butir soal tersebut terbagi dalam lima komponen dengan pengelompokan antar komponen kurang jelas sehingga secara visualisasi akan memberi gambaran pengelompokan yang kabur.

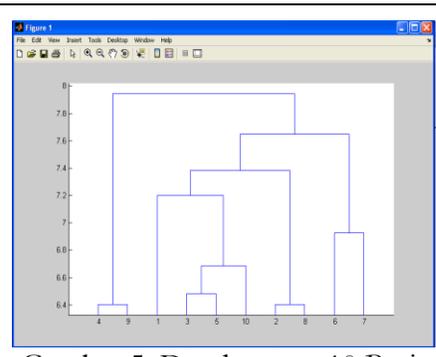


Gambar 3. Plot 10 Butir

Analisis *Fuzzy* C-Means dengan menggunakan program Matlab yakni: `>> X = xlsread('D:\butir soal10_100'); >> x=X'; >> y=pdist(x); >> Z=linkage(y,'single'); >> H=dendrogram(Z)`. Gambar 4 dengan menggunakan metode *linkage single* menunjukkan butir soal (4, 9), butir soal (3, 5) serta butir soal (2, 8) masing-masing membentuk satu komponen sedangkan butir soal 1 mendekati ke butir soal (3,5) membentuk (1,(3,5)) dan butir soal 10 mendekati ke butir soal gabungan butir soal (4,9) dan (1,(3,5)) sedangkan butir soal 7 dan 6 membuat komponen sendiri.



Gambar 4. Dendrogram 10 Butir soal dengan Metode *Linkage Single*

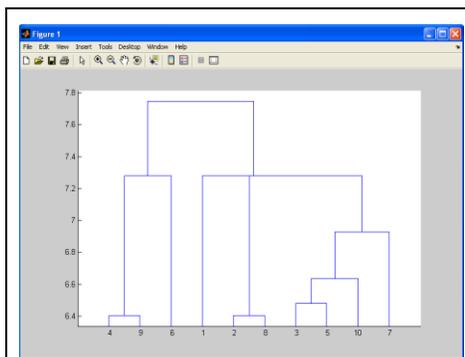


Gambar 5. Dendrogram 10 Butir soal dengan Metode *Linkage Ward*

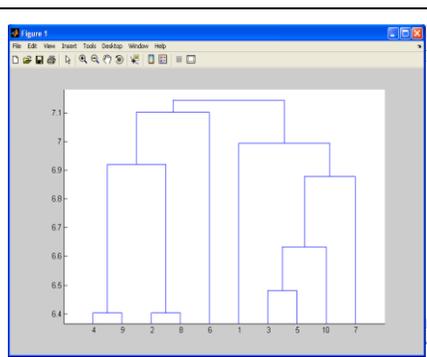
Gambar 5 menggunakan metode *linkage ward* menunjukkan butir soal (4, 9), butir soal (3, 5) serta butir soal (2, 8) dan butir soal (6, 7) masing-masing membentuk satu komponen sedangkan butir soal 10 mendekati ke (3, 5) membentuk komponen sekunder dan butir soal 1 mendekati ke (10, (3, 5)) sekunder dengan menggunakan *linkage ward*. Bila menggunakan metode *linkage ward*, pengelompokan butir soal-butir soal lebih nampak daripada menggunakan metode *single*.

Gambar 6 menggunakan metode *linkage ward complete* menunjukkan butir soal (4, 9), butir soal (3, 5) serta butir soal (2, 8) masing-masing membentuk satu komponen sedangkan butir soal 10 mendekati ke butir soal (3, 5) membentuk (10,(3, 5)), butir soal 6 mendekati ke (4, 9) membentuk komponen (6,(4, 9)), butir soal 1 mendekati ke butir soal (2, 8) membentuk komponen (1,(2, 8)) dan butir soal 7 mendekati ke butir soal gabungan butir soal (10,(3, 5)) dengan menggunakan *linkage complete*.

Gambar 7 menggunakan metode *linkage 'average'* menunjukkan butir soal (4, 9), butir soal (3, 5) serta butir soal (2, 8) masing-masing membentuk satu komponen sedangkan butir soal 10 mendekati ke butir soal (3, 5) membentuk komponen (10,(3, 5)) dan butir soal 7 mendekati ke komponen (10,(3, 5)) membentuk komponen (7,(10,(3, 5))), sedangkan butir soal 6 mendekati ke komponen gabungan (4, 9) dan (2, 8) membentuk komponen (6,(4, 9),(2, 8)) menggunakan metode *linkage average*.



Gambar 6. Dendrogram 10 Butir soal dengan Metode *Linkage Complete*



Gambar 7. Dendrogram 10 Butir soal dengan Metode *Linkage Average*

Hasil simulasi menunjukkan bahwa analisis faktor menghasilkan lima komponen dari 10 butir soal, sedangkan analisis *fuzzy c-means* dengan metode yang berbeda menunjukkan beberapa butir soal konsisten mengelompok dalam tiga komponen yakni (4,9), (2,8) dan (3,5), sedangkan empat butir soal sisanya yakni 1, 6, 7, dan 10 bergerak antar komponen tersebut. Perbedaan banyak komponen tersebut tidak terlalu riskan pada analisis *fuzzy c-means* karena bentuk dendrogram antar butir soal lebih luwes dalam mengaitkan antara butir soal sesuai dengan dasar teori yang digunakan. Misalnya pada Gambar 6, butir soal 10 bisa dimasukkan ke dalam komponen (3, 5) menjadi (10,(3, 5)) dan kalau diperluas lagi, butir soal 7 bisa dimasukkan pada komponen (10,(3, 5)) menjadi komponen (7,(10(3, 5))).

Simulasi I: Data Politomus 3 Butir Soal dan 30 Peserta Tes menggunakan *CFA* dan *fuzzy c means*

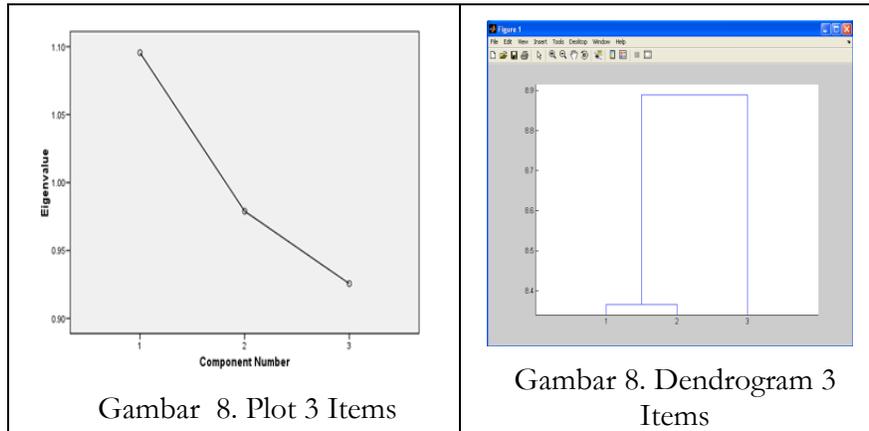
Tabel 5 menunjukkan tingkat kecukupan ukuran sampel sebesar 51% termasuk cukup dengan banyak peserta tes sebesar 30, sedangkan kecocokan data dengan model menggunakan Tes Batlley lewat pendekatan Chi Kuadrat terlihat data berdistribusi normal dimana $\text{sig}=0,977$ lebih besar daripada 0,05.

Tabel 5. KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		,514
Bartlett's Test of Sphericity	Approx. Chi-Square	,202
	df	3
	Sig.	,977

Gambar 8 menunjukkan satu butir soal yang mempunyai nilai eigen yang lebih besar 1 sedangkan dua butir soal lainnya berada di bawah 1 atau sama dengan 1 nilai eigennya. Satu butir soal mempunyai nilai eigen lebih besar daripada 1 maka secara teori hal tersebut tidak dapat dirotasikan.

Analisis *Fuzzy C-Means* Menggunakan Program Matlab yakni: `>>X = xlsread('D:\butir soal3_30'); >>x=X'; >> y=pdist(x); >> Z=linkage(y,'average'); >> H=dendrogram(Z).`



Gambar 9 menunjukkan butir soal 1 dan butir soal 2 berada dalam satu komponen dengan menggunakan metode *linkage average*, sedangkan butir soal 3 berada pada komponen lain. Kecenderungan pola tersebut terjadi pada metode *linkage* yang lain.

Berdasarkan analisis tersebut dengan melibatkan 3 butir soal dengan ukuran sampel 30, kedua metode mempunyai pola yang sama dalam membuat kategori komponen. Walaupun informasi yang diberikan pada analisis faktor lebih banyak, misalnya nilai kecukupan ukuran sampel dan normalisasi serta variansi yang dapat dijelaskan namun hal tersebut tidak terlalu signifikan digunakan terutama dalam kajian eksploratori. Jadi baik menggunakan data dikotomi maupun politomi dengan banyak butir soal 3 dengan ukuran sampel 30 kedua metode sama dalam membuat klasifikasi.

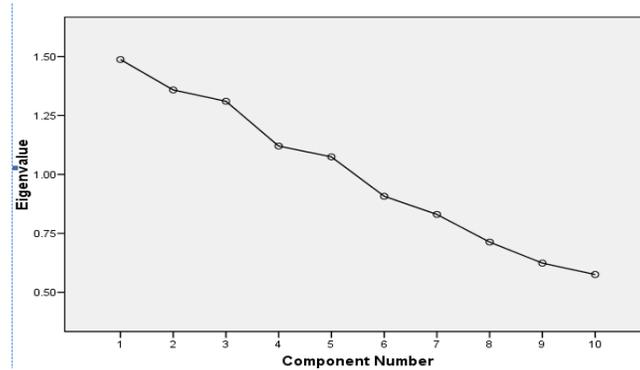
Simulasi II: 3 Butir Soal dan 30 Peserta Tes Menggunakan *CFA* dan *Fuzzy C Means*

Tabel 6 menunjukkan tingkat kecukupan ukuran sampel sebesar 46% termasuk rendah dengan ukuran sampel peserta tes sebesar 100, sedangkan kecocokan data dengan model menggunakan Tes Bartley lewat pendekatan Chi Kuadrat terlihat data berdistribusi normal dimana sig = 0,421 lebih besar daripada 0,05.

Tabel 6. KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		,481
Bartlett's Test of Sphericity	Approx. Chi-Square	46,234
	df	45
	Sig.	,421

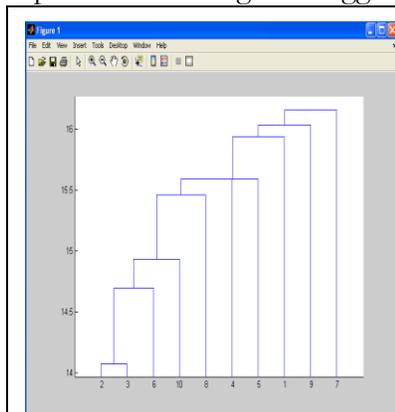
Gambar 10 menunjukkan ada lima butir soal yang mempunyai nilai eigen yang lebih besar 1 sedangkan lima butir soal berada di bawah 1 atau sama dengan 1.



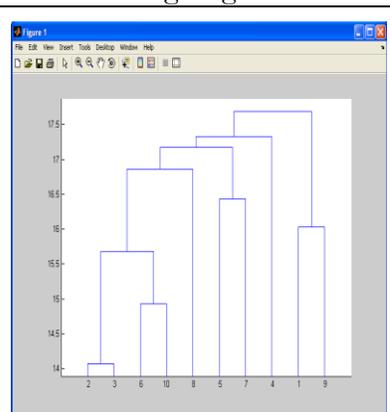
Gambar 10. Plot Sepuluh Butir soal

Analisis *Fuzzy C-Means* dengan Menggunakan Program Matlab yaitu:
`>>X = xlsread('D:\butir soal10_100'); >>x=X'; >> y=pdist(x); >> Z=linkage(y,'single'); >> H=dendrogram(Z).`

Gambar 11 menunjukkan sepuluh butir soal tes terurai secara berjenjang, komponen (2, 3) cakupannya lebih sempit kemudian bergerak saling *linked* dengan butir soal lain sampai pada butir soal 7. Artinya dengan pola tersebut, butir soal-butir soal yang membangun tes mengukur satu kemampuan umum dengan menggunakan metode *linkage single*.



Gambar 11. Dendrogram 10 Butir soal Menggunakan Metode *Linkage Single*

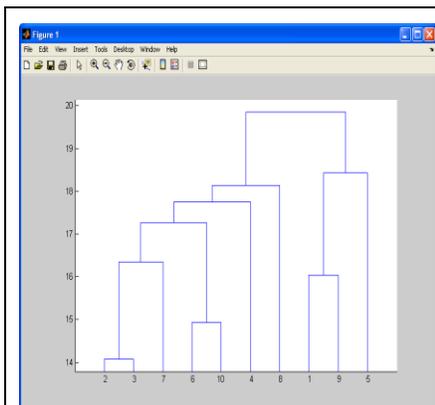


Gambar 12. Dendrogram 10 Butir soal Menggunakan Metode *Linkage Average*

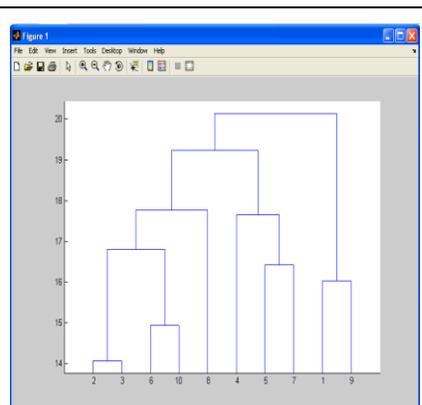
>> $Z = \text{linkage}(y, 'average')$ pada gambar 12 menunjukkan komponen (2, 3), (6, 10), (5, 7) serta (1, 9) masing-masing membentuk komponen terpisah sedangkan butir soal 8 mendekati ke komponen (2, 3) dan (6, 10), butir soal 4 lebih cenderung bergabung ke komponen (8, (2, 3), (6, 10)) dengan komponen (5, 7).

>> $Z = \text{linkage}(y, 'complete')$ pada gambar 13 menunjukkan komponen (2, 3), (6, 10) dan (1, 9) masing-masing membentuk satu komponen sedangkan butir soal 7 mendekati ke komponen (2, 3) membentuk komponen (7, (2, 3)) dan butir soal 5 mendekati ke komponen (1, 9) membentuk komponen (5, (1, 9)), sedangkan butir soal 4 mendekati ke komponen gabungan (7, (2, 3)) dan (6, 10) membentuk komponen (4, (7, (2, 3)), (6, 10)), dan butir soal 8 mendekati ke komponen (4, (7, (2, 3)), (6, 10)) membentuk komponen (7, (4, (7, (2, 3)), (6, 10))) dengan menggunakan metode *linkage complete*.

>> $Z = \text{linkage}(y, 'ward')$ pada gambar 14 menunjukkan komponen (2, 3), (6, 10), (5, 7) dan (1, 9) masing-masing membentuk satu komponen sedangkan butir soal 8 mendekati ke komponen (2, 3) dan (6, 10), dan butir soal 4 mendekati ke komponen (5, 7) membentuk komponen (4, (5, 7)) menggunakan metode *linkage ward*.



Gambar 13. Dendrogram 10 Butir soal Metode *Linkage Complete*



Gambar 14. Dendrogram 10 Butir soal Metode *Linkage Ward*

Berdasarkan hasil analisis *fuzzy c-means* menggunakan keempat metode *linkage* menunjukkan perbedaan cara pengelompokan butir soal, tidak seperti pada data dikotomi ada beberapa butir soal yang konsisten antar metode. Hal tersebut terjadi karena perbedaan dasar pengelompokan yang dikaitkan dengan tingkat variasi jawaban yang ada. Namun demikian, semua metode tersebut dapat ditafsirkan sesuai dengan teori yang melandasi sehingga metode *fuzzy c-means* lebih fleksibel baik dalam penafsiran terhadap pola gambaran menggunakan dendrogram maupun pemilihan metode yang akan digunakan. Artinya penafsiran pola jawaban tetap memperhatikan metode yang digunakan. Walaupun pada penelitian ini lebih cenderung memilih metode *linkage single* untuk data dikotomi namun aturan tersebut tidak bisa dipaksakan terutama data politomous.

Simpulan

Berdasarkan hasil penelitian, dapat disimpulkan bahwa:

1. Metode *fuzzy c-means* lebih luwes dalam menafsirkan pola kecenderungan butir soal terutama butir soal-butir soal yang terjadi secara *linked* menggunakan metode *single* jika dibandingkan dengan analisis faktor.
2. Perbedaan hasil metode *fuzzy c-means* antar metode terutama pada data politomi disebabkan pengaruh lebarnya variasi antara pilihan sehingga kecenderungan perbedaan pengelompokan butir soal sangat menyolok jika dibandingkan dengan data dikotomi.
3. Metode *fuzzy c-means* lebih cocok dilakukan pada tahapan analisis eksplorasi awal dalam mengkaji pengklasifikasian dengan tetap memperhatikan metode yang digunakan.

Saran

Berdasarkan hasil penelitian tersebut, dapat disarankan sebagai berikut:

1. Metode *fuzzy c-means* lebih luwes dalam menginformasikan pengklasifikasian butir soal-butir soal terutama data dikotomi

menggunakan data simulasi. Oleh karena itu, disarankan menggunakan data riil di lapangan.

2. Analisis data tidak melakukan pengajian mendalam perbedaan antara kedua program yang digunakan yakni SPSS dan Matlab terutama analisis algoritma dan koding. Agar interpretasi hasil lebih bermakna disarankan memperhatikan hal tersebut.
3. Metode *fuzzy c-means* mengalami keterbatasan penafsiran jika banyak butir soal lebih besar daripada 30 ke atas sehingga disarankan menggunakan pada tes yang banyak butir soal kecil.

Daftar Pustaka

- Chong Ho Yu, Sharon Osborn Popp, Samuel DiGangi, & Angel Jannasch-Pennell. (2007). *Unidimensionality: A comparison of Rasch modeling, parallel analysis, and TETRA practical assessment*, Research & Evaluation Volume 12, Number 14. Didownload 15 September 2009 dari <http://pareonline.net/pdf/v12n14.pdf>
- Dunn-Rankin, P., Knezek, G.,A., Wallace, S., & Shugian Zhang. (2004). *Scaling methods*. Lawrence Erlbaum Associates Publisher, London.
- Ghozali, Iman & Fuad. (2008). *Structural equation modeling teori, konsep dan aplikasi dengan program liserel 8.80*. Badan Penerbit Universitas Diponegoro, Semarang
- Hambleton, R. K., Swaminathan, H., & Rogers, H.J. (1991). *Fundamentals of item response theory*. Newbury Park, CA: Sage.
- Ledesma, R., D., and Valero-Mora, P. (2007). Determining the number of factors to retain in EFA: an easy-to-use computer program for carrying out Parallel Analysis. *Practical Assessment, Research & Evaluation*, Vol 12, No 2. Didownload 5 September 2009 dari <http://pareonline.net/pdf/v12n2.pdf>

- Lelli, S. (2002). Factor Analysis Fuzzy Sets Theory: Assessing the influence of different techniques on sen's functioning approach. Didownload 17 September 2009 dari <http://www.econ.kuleuven.be/ew/academic/econover/Papers/DPS0121.pdf>
- Moore, A, W. (2001). K-means and hierarchical *klaster* ing. Didownload 14 September 2009 dari <http://www.autonlab.org/tutorials/kmeans11.pdf>.
- Santoso, Agus. (2009). *Computerized adaptive testing untuk pengukuran hasil belajar mahasiswa universitas terbuka*. Disertasi doktor, tidak diterbitkan. Universitas Negeri Yogyakarta, Yogyakarta.
- Sen-Chi Yu and Yuan-Horng Lin. (2008). Applications of fuzzy theory on health care: An example of depression disorder classification based on FCM. WSEAS Transactions on Information science and applications. Issue 1, Volume 5, January. Didownload 20 November 2009 dari <http://www.wseas.us/e-library/transactions/information/2008/24-134.pdf>.
- Suncoko, Cipto. (2009). *Model kontrak partisipasi orang tua pada program paket B setara SMP di wilayah DIY*. Disertasi doktor, tidak diterbitkan. Universitas Negeri Yogyakarta, Yogyakarta.
- Susongko, Purwo. (2009). *Perbandingan keefektifan bentuk tes uraian dan testlet dengan penerepana graded response model (grm)*. Disertasi doktor, tidak diterbitkan. Universitas Negeri Yogyakarta, Yogyakarta.
- Teknomo, K. (2007). K-Means *Klaster* ing Tutorial. Didownload 27 Agustus 2009 dari <http://people.revoledu.com/kardi/tutorial/kMean>
- Yuan H. Li & Schafer, W., D. (2003). *Multidimensional Computerized Adaptive Testing in Recovering Reading and Mathematics Abilities*, Paper presented at the annual meeting of the American Educational Research

Association, Chicago, IL. Didownload tanggal 12 Agustus 2009 dari <http://www.psych.umn.edu/psylabs/catcentral/pdf%20files/li03-01.pdf>.

Wu, M., & Adams, R. (2006). *Modeling mathematics problem solving item response using a multidimensional IRT model*, Mathematics Educational Research Jurnal, Vol. 18 No. 2 p. 93-113. Didownload 27 Agustus 2009 dari http://www.merga.net.au/documents/MERJ_18_2_Wu.pdf

Zimmerman, H. J. 1993. *Fuzzy Set Theory and Its Applications*. 2nd ed. Boston: Kluwer-Nijhoff.

Zwick, W. R., & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological bulletin*, 99, 432-442. Didownload 27 Agustus 2009 dari http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/2e/bc/8f.pdf