

PENGELOMPOKAN PRESTASI MATEMATIKA SISWA INDONESIA BERDASARKAN HASIL SURVEY TIMSS MENGUNAKAN ANALISIS LOGISTIK KELAS LATEN

*Riswan**

Abstract ;

Conventional methods of clustering become weak when meet measured objects with qualitative or categorical data. Latent class logistic analysis can be an alternative method of clustering to overcome this problem. This research is aim to see the application of latent class logistic analysis to cluster the measured objects with qualitative and quantitative variable and at once to find out backgrounds of the clusters. The objects in this research are 2171 eight grade students from 133 schools in Indonesia. There are two results in this research; first in clustering and second in logistic analysis. In clustering, the students have been clustered into four ideal clusters, e.g. 39.16 percent students were in cluster 1, 32.42 percent in cluster 2, 21.46 percent in cluster 3, and 6.97 percent in cluster 4. Each cluster represents the students with very low, low, medium, and high ability in mathematics. In logistic analysis, overall, each cluster has been explained well by covariates e.g. student's interest, attitude, aptitude and motivation on mathematics, parent's social-economic condition, parent's highest education level, teacher's highest education level, teacher's major study of mathematics and educations, teacher's perceptions on schools, school's facilities, etc.

Key Words : Latent class logistic analysis, Covariate, EM algorithm, and Local independence.

A. PENDAHULUAN

Analisis kluster merupakan suatu metode pengelompokan satuan objek pengamatan menjadi beberapa kelompok objek pengamatan berdasarkan peubah-peubah yang dimiliki sehingga objek-objek yang terletak dalam kelompok yang sama relatif lebih homogen dibandingkan dengan objek-objek pada kelompok yang berbeda. Metode pengelompokan klasik seperti metode berhirarki (*single linkage, complete linkage, average linkage, centroid*) dan metode tak

*Dosen Tetap Jurusan Tarbiyah STAIN Samarinda

berhierarchy (*k-means, k-medoid*) sering terkendala masalah data pengukuran yang bersifat kualitatif atau kategorik.

Metode kelas laten dapat dijadikan sebagai metode alternatif dalam mengelompokkan objek pengamatan yang bersifat kategorik. Metode kelas laten adalah suatu metode statistik untuk mengidentifikasi keanggotaan kelas yang tidak terukur (*laten*) antara subjek dengan peubah yang diamati. Metode ini menggunakan fungsi peluang posterior sebagai basis pengelompokannya yang diduga menggunakan metode kemungkinan maksimum (*maximum likelihood*).¹ Metode ini menggambarkan hubungan antara suatu himpunan peubah pengamatan (*manifest/symptom/indikator*) dengan yang tidak diamati yang disebut dengan *peubah laten*. Pengelompokan data pengamatan berdasarkan *indikator* yang melibatkan *kovariat* sebagai peubah penjelas, dapat dikaji lebih jauh dengan menggunakan *analisis logistik kelas laten* sebagai pengembangan dari analisis kelas laten.

Asumsi yang harus dipenuhi dalam analisis logistik kelas laten adalah kebebasan lokal (*local independence*) antar peubah pengamatan. Adanya pelanggaran terhadap asumsi ini akan berpengaruh terhadap tingkat kecocokan model. Kriteria yang digunakan untuk menguji kebebasan lokal adalah nilai *Bivariate Residual* (BVR). Sedangkan kriteria kecocokan model secara keseluruhan digunakan nilai *Bayesian Information Criteria* (BIC) atau *Akaike's Information Criteria* (AIC).

Model kelas laten pertama kali diperkenalkan oleh Lazarsfeld dan Henry² untuk peubah *dichotomous* dan dikembangkan oleh Goodman³ untuk peubah nominal. Akhir-akhir ini model kelas laten telah diperluas lagi untuk peubah campuran (*mixture variable*) baik nominal, ordinal, maupun kontinu. Nainggolan menggunakan analisis kelas laten untuk mengelompokkan pasien demam dengue (DD) dan demam berdarah dengue (DBD).⁴ Hasilnya

¹Vermunt JK, Magidson J. *Latent Class Cluster Analysis*. (Cambridge University Press: Hagenars JA & McCutcheon AL (eds.) *Applied Latent Class Analysis*, 2002), hal. 6-7

²Lazarsfelds P. F., Henry NW. *Latent Structure Analysis*. (Boston: Houghton Mifflin, 1968), hal. 275

³Goodman, L. A. *Exploratory Latent Structure Analysis Using Both Identifiable and Unidentifiable Models*. (Biometrika, 1974), hal. 215-231.

⁴Nainggolan BMH. *Perbandingan Analisis Laten Kelas dengan Kriteria WHO untuk Penggrombolan Pasien Demam Dengue (DD) dan Demam Berdarah Dengue (DBD)* [tesis]. (Bogor: Fakultas Matematika dan Ilmu Pengetahuan Alam, Intstitut Pertanian Bogor, 2009), hal. 1-4

menunjukkan bahwa pengelompokan dengan menggunakan analisis kelas laten mirip dengan pengelompokan menggunakan kriteria dari WHO. Dalam penelitiannya, Nainggolan tidak sampai pada penelusuran terhadap faktor-faktor yang mempengaruhi pasien terkena DD atau DBD.

Sedangkan model logistik kelas laten pertama kali diperkenalkan oleh Chung, Flaherty dan Schafer.⁵ Berdasarkan hasil penelitiannya terhadap siswa SMA dari tahun 1977 s.d 2001, hasilnya menunjukkan bahwa pemakaian marijuana dan sikap-sikap siswa terhadap kondisi moral dan sosialnya dapat disimpulkan dengan baik ke dalam empat model kelas laten. Penerapan model logistik multinomial terhadap respon laten tersebut menunjukkan bahwa pengelompokan siswa dalam kelas laten sangat berkaitan dengan faktor-faktor demografis, gaya hidup, *political beliefs*, dan agama.

Tulisan ini bertujuan untuk mengkaji dan menerapkan model analisis kelas laten dalam mengelompokkan objek pengamatan yang bersifat kategorik sekaligus mencari faktor-faktor yang melatarbelakanginya. Dalam tulisan ini analisis logistik kelas laten diterapkan untuk mengelompokkan prestasi siswa Indonesia berdasarkan hasil survey TIMSS (*Trend in International Mathematics and Science Study*) yang dilakukan oleh lembaga IEA (*International Association for the Evaluation of Educational Achievement*). Setelah dilakukan pengelompokan kemudian dicari faktor-faktor yang melatarbelakangi prestasi siswa tersebut pada tiap kelompoknya.

Sebagaimana diketahui bahwa lembaga IEA melakukan survey secara berkala setiap empat tahun sekali, yang dilakukan terhadap siswa, guru, dan sekolah di sejumlah negara di dunia. Indonesia telah mengikuti proyek ini sejak tahun 1995, hasil survey terakhir, yaitu tahun 2007, menunjukkan bahwa Indonesia menempati ranking ke-35 dari 49 negara peserta. Ranking ini tidak jauh berbeda dengan hasil survey tahun 2003 dan 1999, di mana prestasi siswa Indonesia di tingkat Asia Tenggara saja masih di bawah Singapura, Malaysia, bahkan Thailand. Sehingga menarik untuk diteliti apa yang melatarbelakangi rendahnya prestasi matematika siswa Indonesia tersebut.

⁵Chung H. et al., *Latent Class Logistic Regression: Application to marijuana use and attitudes among high school seniors*. (Journal of The Royal Statistical Society, 2006), hal. 723-743.

Kajian yang berkaitan dengan hasil survey TIMSS telah dilakukan di antaranya oleh Santoso.⁶ Hasilnya menunjukkan bahwa, secara umum, faktor-faktor seperti sikap atau motivasi belajar matematika siswa, persepsi siswa terhadap sekolah, persepsi siswa terhadap matematika, minat belajar siswa, perilaku siswa di sekolah, keadaan sosial ekonomi orang tua, latar belakang guru, penilaian guru terhadap sekolah, serta sarana dan prasarana sekolah sangat berpengaruh terhadap prestasi matematika siswa. Namun demikian, belum ada riset yang mencoba mengelompokkan prestasi matematika siswa tersebut sekaligus mencari faktor-faktor yang melatarbelakanginya. Oleh karena itu penulis tertarik untuk melakukan kajian ini dengan menggunakan *analisislogistik kelas laten*.

B. TINJAUAN TENTANG METODOLOGI KAJIAN

Bahan dalam penelitian ini adalah data sekunder hasil survey TIMSS tahun 2007 melalui Pusat Penelitian Pendidikan (Puspendik) Badan Penelitian dan Pengembangan Departemen Pendidikan Nasional. Survey dilakukan melalui angket kuesioner yang ditujukan kepada siswa, guru dan sekolah di sejumlah daerah di Indonesia. Data berasal dari 2171 siswa SLTP kelas 8 dari 133 sekolah baik negeri maupun swasta. Data berkaitan dengan kemampuan matematika siswa (aljabar, data dan peluang, bilangan, dan geometri) dan data latar belakang siswa, guru dan sekolah. Data tersebut dikelompokkan ke dalam dua peubah yaitu peubah indikator berupa skor matematika siswa, dan peubah kovariat berupa latar belakang siswa, guru dan sekolah.⁷

Tahapan yang dilakukan dalam kajian ini adalah:

1. Menyiapkan data dengan indikator y dan kovariat x .
2. Membentuk kelas laten berdasarkan nilai peluang dari semua sampel menggunakan *software Latent Gold 4.0*. Dalam hal ini, peubah y_1 sampai dengan y_4 masing-masing berupa peubah nominal dimana peubah y_i berbentuk politomus dengan lima kategori (sangat rendah, rendah, sedang, tinggi, *advance*). Sebaran multinomial berganda y_i dengan peubah kovariat x_i memiliki fungsi sebaran peluang sebagai ;

⁶Santoso A. *Faktor-faktor yang mempengaruhi prestasi siswa: tinjauan berdasarkan data TIMSS 2007* (Jakarta: Badan Penelitian dan Pengembangan Departemen Pendidikan Nasional, 2009), hal. 4-6

⁷Sumber data: Pusat data TIMSS, *diekstrak* menggunakan software IDBAnalyzer.

$$f(y_h | x_i, \alpha_k) = \sum_{k=1}^K \pi_k(x_i) g(y_i | \alpha_k) = \sum_{k=1}^K \pi_k(x_i) \prod_{i=1}^P \prod_{s=1}^{S_p} (\alpha_{k(is)})^{y_{h(is)}}$$

Dimana M adalah jumlah parameter, $M = (K - 1) + K * (\sum_i S_i - 1)$, $DF(K) = \prod_{i=1}^P S_i - M - 1$, dan $\pi = (\pi_1, \pi_2, \dots, \pi_K)$ adalah vektor proporsi campuran dari K kelas laten, p adalah jumlah peubah yaitu 4, dan S banyaknya kategori yaitu 5.

3. Pendugaan parameter menggunakan algoritma EM, yaitu:

a. Definisikan nilai awal $\theta^{(0)} = (\pi^{(0)}, \alpha^{(0)}, \mu^{(0)}, \sigma^{(0)})$.

b. Hitung sebaran peluang bersama:

$$f(y_h | x_i, \hat{\alpha}_k) = \sum_{i=1}^K \hat{\pi}_k(x_i) g(y_h | \hat{\alpha}_k)$$

sebaran multinomial dengan $p = 4$, dan $S = 5$ adalah

$$f(y_h | x_i, \alpha_k) = \sum_{k=1}^K \pi_k(x_i) g(y_i | \alpha_k) = \sum_{k=1}^K \pi_k(x_i) \prod_{i=1}^4 \prod_{s=1}^5 (\alpha_{k(is)})^{y_{h(is)}}$$

c. Tahapan E, hitung $(\hat{P}(k | y_h))^{(r)}$, $h = 1, \dots, 2171$ dan $k = 1, \dots, 6$

$$\hat{P}(k | y_h)^{(r)} = \frac{\hat{\pi}^{(r)} g(y_h, \hat{\alpha}_k)^{(r)}}{\sum_{i=1}^K \hat{\pi}_i^{(r)} g(y_h, \hat{\alpha}_i)^{(r)}}$$

menyatakan peluang bersyarat yang menyatakan y_h muncul dari k.

d. Tahapan M, sesuaikan pendugaan parameter yang baru yaitu

$$\hat{\pi}_k = \frac{\sum_{h=1}^n \hat{P}(k | y_h)}{n}$$

untuk peubah **nominal**, pendugaan peluang bersyarat $y_j = s$ adalah

$$\hat{\alpha}_{k(is)} = \frac{y_{ih(s)} \sum_{h=1}^n \hat{P}(k | y_h)}{n \hat{\pi}_k}$$

e. Ulangi tahap 2 dan 3 sampai konvergen.

4. Ketika algoritma EM mencapai optimal, dilanjutkan ke metode Newton Raphson.
5. Memilih klaster terbaik dengan menggunakan nilai BIC.
6. Memeriksa asumsi kebebasan lokal dengan menggunakan nilai BVR.
7. Menguji signifikansi nilai dugaan parameter dengan *Wald Chi-Square Statistic*.
8. Interpretasi hasil analisis logistik kelas laten.
9. Simpulan dan saran

C. KONSEP DASAR ANALISIS KELAS LATEN

Secara umum model kelas laten dinyatakan dengan sebaran gabungan dari peubah-peubah yang diamati pada data yang memiliki berbagai macam skala pengukuran. Misalkan (y_1, y_2, \dots, y_p) dinotasikan sebagai vektor dari p peubah indikator dimana setiap peubah memiliki sebaran bersyarat dalam keluarga eksponensial seperti Bernoulli, Poisson, multinomial, dan normal. Misalkan y_{ih} adalah nilai dari h ($h=1, 2, \dots, n$) sampel objek untuk peubah ke- i . Vektor baris $y'_h = (y_{1h}, y_{2h}, \dots, y_{ph})$ mengacu ke bentuk respon dari i objek, maka model kelas laten didefinisikan sebagai sebaran peluang bersama yaitu:

$$f(y_h | \alpha_k) = \sum_{k=1}^K \pi_k g(y_h | \alpha_k)$$

dimana $g(y_h | \alpha_k)$ adalah sebaran y_i dengan parameter model α_k , π_k adalah peluang awal kelas laten pada data y , α_k adalah peluang suatu objek pada kelas k , dengan k banyaknya kelas ($k=1, 2, \dots, K$).

Peubah **biner**: Pada kasus dimana peubah y_i berbentuk biner (0 dan 1), sebaran ini diasumsikan berbentuk sebaran Bernoulli berganda, maka sebaran peluangnya adalah

$$f(y_h | \alpha_k) = \sum_{k=1}^K \pi_k g(y_i | \alpha_k) = \sum_{k=1}^K \pi_k \prod_{i=1}^P \alpha_{ik}^{y_i} (1 - \alpha_{ik})^{1-y_i}$$

dengan α_{ik} adalah peluang suatu objek pada gerombol k , M adalah jumlah parameter, dimana $M = (K - 1) + K^*p$, DF adalah derajat kebebasan, dimana $DF(K) = 2^p - M - 1$.

Peubah **nominal**: Pada kasus dimana peubah y_i berbentuk politomus, maka sebaran peluangnya adalah

$$f(y_h | \alpha_k) = \sum_{k=1}^K \pi_k g(y_i | \alpha_k) = \sum_{k=1}^K \pi_k \prod_{i=1}^P \prod_{s=1}^{S_p} (\alpha_{k(is)})^{y_{h(is)}}$$

dengan M adalah jumlah parameter dimana $M = (K - 1) + K^*(\sum_i S_i - 1)$, $DF(K) = \prod_{i=1}^P S_i - M - 1$, dan $\pi = (\pi_1, \pi_2, \dots, \pi_K)$ adalah vektor proporsi campuran dari K kelas.

Peubah **ordinal**: Pada kasus dimana peubah y_i berbentuk ordinal, maka sebaran peluangnya adalah

$$f(y_h | \alpha_k) = \sum_{k=1}^K \pi_k g(y_i | \alpha_k) = \sum_{k=1}^K \pi_k \prod_{i=1}^P \prod_{s=1}^{L_p} (\gamma_{ik(s)} - \gamma_{ik(s-1)})^{y_{h(i,s)}}$$

dengan $M = (K - 1) + K^*(\sum_i L_i - 1)$, $DF(K) = \prod_{i=1}^p L_i - M - 1$.

Peubah **kuantitatif**: Pada kasus dimana peubah y_i berbentuk kuantitatif, sebarannya diasumsikan normal, maka sebaran peluangnya adalah

$$f(y_h | \mu_k, \sigma_i^2) = \sum_{k=1}^K \pi_k \prod_{i=1}^p (2\pi)^{-\frac{1}{2}} \sigma_i^{-\frac{1}{2}} \exp\left[-\frac{1}{2\sigma_i^2} (y_i - \mu_{ik})^2\right]$$

dengan μ_{ik} adalah lokasi parameter dari peubah kontinu y_i dalam kelas k . σ_i^2 adalah ragam dari peubah ke- i , dan jumlah parameter $(2p + 1) * K - 1$.

D. MODEL LOGISTIK KELAS LATEN

Model kelas laten dengan melibatkan kovariat adalah sebagai berikut:

$$f(y_h | x_i, \alpha_k) = \sum_{k=1}^K \pi_k(x_i) \prod_{i=1}^p g(y_{hi} | \alpha_{ik})$$

dimana x_i dinotasikan sebagai nilai kovariat objek h , maka model logistik kelas laten didefinisikan sebagai berikut:

$$\pi_k(x_i) = \frac{\exp(\beta_k x_i)}{\sum_{k=1}^K \exp(\beta_k x_i)}$$

Parameter model logit diduga secara bersamaan dengan model kelas laten menggunakan algoritma EM. Pengujian nilai signifikansi dugaan parameter dinyatakan dengan rumusan hipotesis berikut:

$H_0 : \beta_i = 0$ (koefisien logit tidak berpengaruh terhadap model)

$H_0 : \exists \beta_i \neq 0$ (ada koefisien logit yang berpengaruh terhadap model)

Dengan statistik uji menggunakan *Wald Chi-Square Statistic* yaitu $W = (\hat{\beta} / SE(\hat{\beta}))^2$.

Jika digunakan tarap nyata α , maka tolak H_0 jika nilai $W > \chi_{p,\alpha}^2$ atau *p-value* $\leq \alpha$.

1. Pendugaan parameter

Dua metode utama untuk menduga parameter pada analisis kelas laten adalah kemungkinan maksimum (EM) dan metode Newton-Raphson. Fungsi likelihood untuk analisis model kelas laten campuran adalah:

$$L = \sum_{h=1}^n \log f(y_h) \text{ dan } L(\theta | y) = \sum_{h=1}^n \log \sum_{k=1}^K \pi_k g(y_h | \alpha_k)$$

Dalam hal ini peubah pengamatan bebas bersyarat pada setiap kluster k . Dinotasikan $x = (x_1, \dots, x_K)$ dengan $x_K = (x_{1K}, \dots, x_{nK})$, $x_{ik} = 1$ jika y_i muncul dari kluster k , $x_{ik} = 0$ untuk lainnya. Vektor indikator yang tidak diketahui dari K kluster memiliki bentuk likelihood lengkap yaitu:

$$L(\theta | y, x) = \sum_{i=1}^n \sum_{k=1}^K x_{ik} \log \pi_k g(y_i | \alpha_k)$$

Log-likelihood tersebut dimaksimumkan dengan menggunakan algoritma EM dengan kendala $\sum_{k=1}^K \pi_k = 1$. Kendala tambahan dipersyaratkan pada kasus nominal dan ordinal, yaitu $\sum_{s=1}^{S_i} \alpha_{k(is)} = 1$ dan $\sum_{s=1}^{L_i} \alpha_{k(is)} = 1$.

2. Ukuran kecocokan model

Kriteria pememilihan kluster atau model terbaik dapat digunakan beberapa kriteria seperti: Statistik *Chi-Square*, *Bayesian Information Criteria* (BIC), dan *Akaike's Information Criteria* (AIC). Model terbaik adalah model dengan nilai AIC dan BIC terkecil. (Vermunt & Magidson 2001).

$$AIC = -2[\max \ln(L)] + 2m \text{ dan } BIC = -2[\max \ln(L)] + 2m \log(n)$$

dalam hal ini, m adalah banyaknya parameter, n adalah ukuran sampel, dan L adalah fungsi kemungkinan likelihood.

3. Asumsi kebebasan lokal

Analisis kelas laten mensyaratkan antar peubah harus saling bebas pada suatu kelas laten tertentu yang disebut dengan kebebasan lokal. Pelanggaran terhadap asumsi kebebasan lokal dilihat dari nilai *Bivariate Residual* (BVR) yaitu nilai *Pearson Chi-Square* dibagi dengan derajat bebasnya, yaitu:

$$X^2 = \sum_{i=1}^j \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \text{ dan } BVR = \frac{X^2}{df}$$

dengan O_{ij} adalah frekuensi observasi, E_{ij} frekuensi harapan, dan $df = (p-1)(k-1)$. Asumsi kebebasan lokal dipenuhi jika nilai $BVR < 3.84$.

4. Prestasi matematika siswa

Prestasi siswa dalam bidang matematika bisa dilihat dari kemampuan mereka dalam memecahkan soal-soal atau permasalahan matematika. Bidang-bidang matematika yang dipelajari siswa di sekolah bisa digolongkan ke dalam empat kategori yaitu aljabar, data dan peluang, bilangan, dan geometri. TIMSS mengelompokkan kemampuan matematika baik aljabar, data dan peluang, bilangan maupun geometri berdasarkan *Math International Benchmark* yaitu: kemampuan siswa dengan skor kurang dari 400, antara 400 sampai kurang dari 475, antara 475 sampai kurang dari 550, antara 550 sampai kurang dari 625, dan skor 625 ke atas. Berdasarkan hasil survey 2007, skor matematika siswa Indonesia berada pada rata-rata 397,1 termasuk ke dalam kategori sangat rendah.

5. Faktor-faktor yang Mempengaruhi Prestasi Siswa

Secara umum banyak sekali faktor yang dapat mempengaruhi prestasi belajar siswa baik internal maupun eksternal. Faktor internal berasal dari dalam diri siswa seperti aspek fisiologis dan psikologis. Aspek psikologis dapat mempengaruhi kuantitas dan kualitas perolehan pembelajaran siswa, beberapa hal yang dipandang penting adalah tingkat kecerdasan, sikap siswa terhadap pelajaran, bakat, minat dan motivasi siswa.⁸ Sedangkan pendidikan orang tua, cita-cita pendidikan siswa, jumlah buku yang dimiliki siswa di rumah, ketersediaan perangkat komputer, sosial ekonomi, waktu pengerjakaan pekerjaan rumah merupakan faktor eksternal dari siswa yang berpengaruh terhadap prestasi akademiknya.⁹

Khusus dalam bidang matematika, Santoso telah merangkum faktor-faktor baik internal maupun eksternal yang dapat mempengaruhi prestasi matematika siswa yaitu: sikap atau motivasi belajar matematika siswa, persepsi siswa terhadap sekolah, persepsi siswa terhadap matematika, minat belajar siswa, perilaku siswa di sekolah, keadaan sosial ekonomi orang tua, latar

⁸Syah M. *Psikologi Pendidikan dengan Pendekatan Baru (Edisi Revisi)*. (Bandung: PT Remaja Rosda Karya, 2005), hal. 144

⁹Mullis IV S *et al.* *TIMSS 2007: Assessment Frameworks from IEA's Trends in International Mathematics and Science Study at the fourth and Eight Grades*. (Boston College: TIMSS International Study Center, 2005), hal. 35-37

belakang guru, penilaian guru terhadap sekolah, serta sarana dan prasarana sekolah.¹⁰

6. Variabel penelitian

Variabel atau peubah penelitian terdiri dari peubah *indikator* dan *kovariat* atau peubah penjelas yaitu sebagai berikut:

Indikator:

- y_1 : Kemampuan aljabar
- y_2 : Kemampuan data dan peluang
- y_3 : Kemampuan bilangan
- y_4 : Kemampuan geometri

Kovariat:

- x_1 : Lama mengajar (0-2, 2-5, 5-9, 9-14, 14-20, 20-27, >27 tahun)
- x_2 : Tingkat pendidikan guru (SLTA, D1/D2, D3/D4, S1,S2/S3)
- x_3 : Latar belakang pendidikan guru (linear, agak linear, tidak linear)
- x_4 : Persepsi guru terhadap sekolah (tinggi, sedang, rendah, sgtrendah)
- x_5 : Banyak buku yang dimiliki siswa (0-11, 12-25, 26-100, 100-200, >200 buku)
- x_6 : Sosial ekonomi orang tua (sgt tinggi, tinggi, sedang, rendah, sgt rendah)
- x_7 : Tingkat pendidikan orang tua (sgt rendah, rendah, sedang, tinggi, sgttinggi)
- x_8 : Motivasi belajar siswa (sgttinggi, tinggi, sedang, rendah, sgtrendah)
- x_9 : Persepsi siswa terhadap matematika (jelek, biasa saja, baik)
- x_{10} : Minat belajar matematika siswa (sgttinggi, tinggi, sedang, rendah)
- x_{11} : Persepsi siswa terhadap sekolah (jelek, biasa saja, baik)
- x_{12} : Intensitas pemberian PR/minggu (setiap hari, 3-4 kali, 1-2 kali, kurang dari sekali, tidak pernah)
- x_{13} : Lama waktu mengerjakan PR (0,1-15,16-30,31-60,61- 90,>90 menit)
- x_{14} : Perilaku siswa (baik, sedang, jelek, sangat jelek)
- x_{15} : Penguatan (tinggi, sedang, rendah)
- x_{16} : Sarana prasarana sekolah (sgt. kurang, kurang, sedang, banyak, sgt. banyak)

¹⁰Santoso, *Faktor-faktor yang Mempengaruhi*, hal. 4-6

x_{17} : Penggunaan waktu luang (normal, sedang, jelek, sangat jelek)

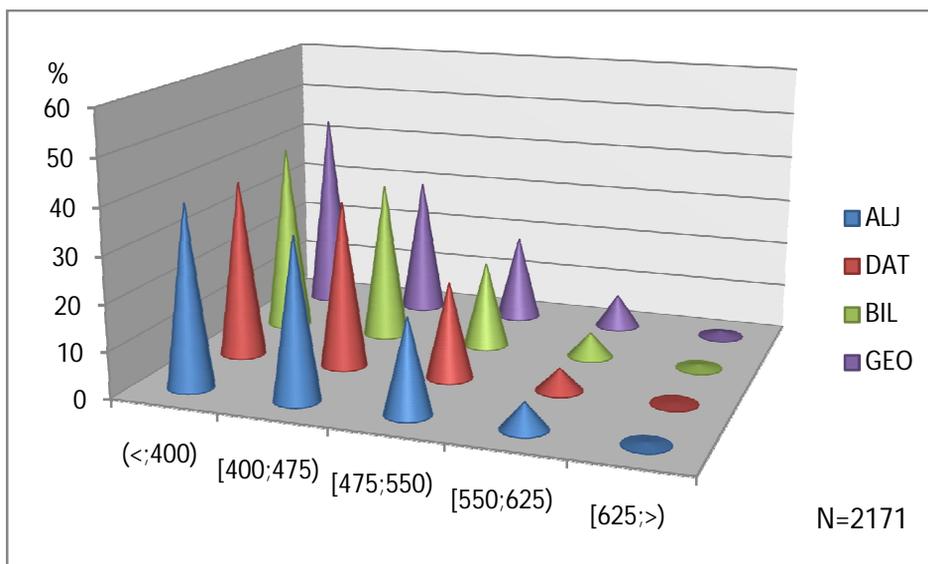
E. HASIL KAJIAN

a. Deskripsi data

Berdasarkan hasil survey TIMSS tahun 2007, diperoleh informasi bahwa prestasi siswa Indonesia khususnya siswa SLTP kelas 8 masih jauh di bawah standar Internasional. Skor matematika siswa Indonesia sebagian besar, yaitu sekitar 40%-nya, masih di bawah 400 baik untuk skor aljabar, data dan peluang, bilangan, maupun geometri, dan hanya sebagian kecil saja diantara mereka yang dapat meraih skor di atas 625, paling banyak 6% untuk skor aljabar, selebihnya jauh di bawah 6%. Kemampuan siswa Indonesia dalam bidang matematika tersebut tergambar dalam Tabel 1 dan gambar 1 berikut ini.

Tabel 1 Skor matematika siswa SLTP kelas 8

SKOR	ALJABAR		DATA & PELUANG		BILANGAN		GEOMETRI	
	Jumlah	%	Jumlah	%	Jumlah	%	Jumlah	%
(<400)	850	39,15	834	38,42	889	40,95	942	43,39
[400;475)	748	34,45	779	35,88	746	34,36	648	29,85
[475;550)	438	20,18	448	20,64	404	18,61	402	18,52
[550;625)	122	5,62	103	4,74	113	5,20	154	7,09
[625;>)	13	0,60	7	0,32	19	0,88	25	1,15
Total	2171	100	2171	100	2171	100	2171	100



Gambar 1 Skor matematika siswa SLTP Kelas 8

b. Hasil analisis kelas laten

1. Pemilihan model terbaik

Pemilihan model terbaik dalam analisis kelas laten mengandung pengertian berapa kelas terbaik untuk mengelompokan objek pengamatan. Berikut ini disajikan berbagai model hasil analisis kelas laten.

Tabel 2 Kecocokan model

		LL	BIC(L L)	N pa r	L ²	df	p- value
Mod el2	2- Kelas	- 8571,1 799	17833, 825	90	17092, 453	208 1	1,3e- 2311
Mod el3	3- Kelas	- 7801,1 951	16862, 393	16 4	15552, 484	200 7	1,9e- 2052
Mod	4-	-	16859,	23	14980,	193	3,4e-

el4	Kelas	7515,2	071	8	624	3	1977
		652					
Mod	5-	-	17095,	31	14648,	185	2,5e-
el5	Kelas	7349,1	4	2	415	9	1947
		607					

Berdasarkan Tabel 2 diketahui bahwa nilai BIC terkecil diperoleh pada Model4, yaitu model dengan 4 kelas dengan *p-value* signifikan. Maka untuk selanjutnya, dipilih model dengan 4 kelas sebagai model terbaik untuk mengelompokkan prestasi siswa Indonesia berdasarkan kemampuannya dalam bidang matematika yaitu ajar, data dan peluang, bilangan, dan geometri.

2. Pemeriksaan asumsi kebebasan lokal

Pemeriksaan asumsi kebebasan lokal dilakukan terhadap model terbaik yang sudah terpilih yaitu model dengan 4 kelas. Berikut ini disajikan nilai *Bivariate Residual* (BVR) sebagai kriteria yang digunakan untuk menilai kondisi kebebasan lokal.

Tabel 3 Nilai *Bivariate Residual* (BVR)

Indikator	y_1	y_2	y_3	y_4
y_1				
y_2	1,9772			
y_3	2,0791	1,9582		
y_4	0,8589	1,2267	1,9658	
Kovariat	y_1	y_2	y_3	y_4
x_{11}	5,6266	0,6723	9,3972	24,9932

Berdasarkan Tabel 3 diperoleh informasi bahwa terdapat gangguan asumsi kebebasan lokal pada peubah kovariat x_{11} yaitu persepsi guru terhadap sekolah dengan peubah indikator y_1, y_2 , dan y_4 dengan nilai BVR > 3.84 , dan jika dilihat pada pada Tabel 4, peubah x_{11} tidak berpengaruh nyata terhadap model dengan *p-value* sebesar 1. Sehingga untuk mengatasi gangguan terhadap asumsi kebebasan lokal tersebut dilakukan dengan mengeluarkan peubah x_{11} tersebut dari model, kemudian dilakukan pemodelan ulang terhadap model dengan 4 kelas tersebut.

Tabel 4 Nilai signifikansi parameter

x_{11}	Kelas1	Kelas2	Kelas3	Kelas4	p -value
Baik	0,566	1,0227	- 1,9949	0,4062	0,1888
Kurang Baik	0,5579	0,9767	- 2,0971	0,5626	
Tidak Baik	- 1,1239	- 1,9994	4,092	- 0,9688	

Setelah dilakukan pemodelan ulang, sebagaimana terlihat pada Tabel 5, diperoleh Model 5 dengan 4 kelas dengan keadaan yang lebih baik dimana nilai BIC-nya lebih kecil dari model sebelumnya, serta tidak ada lagi nilai BVR > 3.84, sehingga model sudah memenuhi asumsi kebebasan lokal.

Tabel 5 Kecocokan model

		LL	BIC(LL)	Npar	L ²	df	p-value
Model4	4- Kelas	- 7515,2652	16859,071	238	14980,624	1933	3,4e- 1977
Model5	4- Kelas	-7516,196	16814,835	232	14982,486	1939	7,0e- 1975

3. Karakteristik model kelas laten berdasarkan peubah indikator

Berdasarkan uraian terdahulu bahwa model kelas laten terbaik adalah model dengan 4 kelas. Karakteristik dari masing-masing kelas tersebut berdasarkan peubah indikatornya dapat dilihat pada Tabel 6.

Tabel 6 Karakteristik model berdasarkan peubah indikator

		Kelas1	Kelas2	Kelas3	Kelas4
	Ukuran kelas				
	Skor				
Indikator	Aljabar (<:400)	0,8251	0,2107	0,0006	0,0006
	[400;475)	0,1672	0,672	0,285	0,0007

(y_1)	[475;550)	0,0077	0,1163	0,6446	0,3257
	[550;625)	0	0,001	0,0698	0,587
	[625;>)	0	0	0	0,0859
Data&Peluang	(<;400)	0,8067	0,1988	0,0177	0,0011
(y_2)	[400;475)	0,1826	0,6319	0,3786	0,0176
	[475;550)	0,0108	0,1655	0,5472	0,446
	[550;625)	0	0,0038	0,0565	0,489
Bilangan	[625;>)	0	0	0	0,0463
	(<;400)	0,9172	0,155	0,0003	0,0007
	[400;475)	0,0827	0,7719	0,2841	0,0006
(y_3)	[475;550)	0,0001	0,073	0,6888	0,2089
	[550;625)	0	0	0,0268	0,6642
	[625;>)	0	0	0	0,1256
Geometri	(<;400)	0,9027	0,2448	0,0048	0,0007
	[400;475)	0,0967	0,6541	0,2225	0,0118
	[475;550)	0,0005	0,0994	0,6709	0,1257
(y_4)	[550;625)	0	0,0017	0,1018	0,6966
	[625;>)	0	0	0	0,1652

Berdasarkan Tabel 6 tersebut dapat digambarkan bahwa kelompok pertama atau kelas1 dengan proporsi paling besar 39,16% atau sekitar 850 siswa adalah kelompok siswa yang prestasinya paling rendah. Keadaan ini dapat dilihat dari sebagian besar siswa yang berada pada kelompok ini memperoleh skor (<;400) dengan proporsi sebesar 82,51% untuk skor aljabar, 80,67% untuk skor data dan peluang, 91,27% untuk skor bilangan, 90,27% untuk skor geometri. Tidak ada sama sekali siswa yang memperoleh skor 550 atau lebih, baik untuk aljabar, data dan peluang, bilangan, maupun geometri.

Kelompok kedua atau kelas 2 dengan proporsi sebesar 32,42% atau sekitar 704 siswa sedikit lebih baik perolehan skor matematikanya dari kelompok yang pertama. Sebagian besar diantara mereka memperoleh skor [400;475) dengan proporsi sebesar 67,2% untuk skor aljabar, 63,16% untuk skor data dan peluang, 77,19% untuk skor bilangan, 65,41% untuk skor geometri, dan sisanya ada yang memperoleh skor (<;400) dan [475;550), dan sama sekali tidak ada siswa yang memperoleh skor [625;>) baik untuk skor aljabar, data dan peluang, bilangan, maupun geometri.

Kelompok ketiga atau kelas3 dengan proporsi sebesar 21,46% atau sekitar 466 siswa. Sebagian besar diantara mereka memperoleh skor [475;550) dengan proporsi sebesar 64,46% untuk skor aljabar, 54,16% untuk skor data dan peluang, 67,09% untuk skor bilangan, 68,88% untuk skor geometri, dan sedikit sisanya ada yang memperoleh skor (\leq ;400), [475;550) dan [550;625). Walaupun siswa di kelompok ini skornya lebih baik dibanding kelompok pertama dan kedua, namun tidak ada siswa yang memperoleh skor [625;>) baik untuk aljabar, data dan peluang, bilangan, maupun geometri.

Kelompok terakhir yaitu kelas 4 dengan proporsi paling kecil yaitu sebesar 6,97% atau hanya sekitar 151 siswa saja adalah kelompok siswa yang prestasinya paling tinggi. Sebagian besar diantara mereka memperoleh skor [550;625) dengan proporsi sebesar 58,7% untuk skor aljabar, 48,9% untuk skor data dan peluang, 66,42% untuk skor bilangan, 69,66% untuk skor geometri. Sebagian lain siswa memperoleh skor [475;550) dengan proporsi sebesar 32,57% untuk skor aljabar, 44,6% untuk skor data dan peluang, 20,89% untuk skor bilangan, 12,57% untuk skor geometri. Walaupun siswa di kelompok empat ini memiliki skor paling tinggi dibandingkan kelompok lainnya, namun sedikit diantara mereka yang memperoleh skor [625;>) yaitu 4,63% untuk skor aljabar, 8,59% untuk skor data dan peluang, 12,56% untuk skor bilangan, 16,52% untuk skor geometri.

Jika kelompok tersebut didefinisikan ke dalam suatu kategori prestasi siswa dalam bidang matematika, maka siswa yang berada pada kelompok pertama termasuk kategori *sangat rendah* (SR), kedua *rendah* (R), ketiga *sedang* (S), dan keempat *tinggi* (T). Tidak ada kategori *sangat tinggi*, hal ini bisa dipahami bahwa prestasi siswa Indonesia masih di bawah rata-rata skor TIMSS yaitu sebesar 500 dan sangat sedikit sekali siswa yang memperoleh skor 625 ke atas.

4. Parameter model logistik kelas laten

Pendugaan parameter logistik kelas laten dilakukan dengan menggunakan algoritma EM. Ketika algoritma EM telah mencapai nilai yang optimum, proses dilanjutkan dengan metode Newton-Raphson. Dalam aplikasinya, kedua tahapan tersebut dilakukan dengan menggunakan *software Latent Gold 4.0*.

Hasil pendugaan parameter pada analisis logistik kelas laten menunjukkan bahwa pada tarap nyata 15%, hampir semua parameter peubah kovariat x memiliki pengaruh nyata terhadap model kecuali x_{15} yaitu peubah penguatan pembelajaran. Hal ini menunjukkan bahwa secara umum latar belakang siswa, guru, dan sekolah sangat berpengaruh terhadap prestasi siswa dalam bidang matematika.

Model logistik kelas latendengan empat kelas ($K=4$), 16 peubah kovariat ($P=16$), dan 2171 pengamatan ($I=2171$) dapat dinyatakan sebagai berikut:

$$\pi_k(x_i) = \frac{\exp(\beta_k x_i)}{\sum_{k=1}^K \exp(\beta_k x_i)} \Leftrightarrow \pi_k(x_i) = \frac{\exp(\beta_k x_i)}{\sum_{k=1}^4 \exp(\beta_k x_i)}$$

dengan $k = 1, 2, 3, K$, $p=1, 2, 3, \dots, P$, dan $i=1, 2, 3, \dots, I$.

Berdasarkan hasil pendugaan parameter logistik kelas laten, maka dapat diduga besaran peluang suatu pengamatan atau individu dengan karakteristik peubah kovariat tertentu masuk ke dalam kelompok yang ada. Dalam hal ini, dapat diperkirakan besarnya peluang seorang siswa dengan latar belakang tertentu masuk ke dalam kelompok siswa yang prestasinya *sangat rendah* (SR), *rendah* (R), *sedang* (S), atau *tinggi* (T). Sebagai contoh, siswa dengan latar belakang yang kurang baik sebagaimana terlihat pada Tabel 7, memiliki peluang yang besar untuk masuk ke dalam kelompok siswa yang prestasinya *sangat rendah* (SR) yaitu sebesar 99,97%, dan sangat kecil sekali peluangnya untuk bisa masuk ke dalam kelompok siswa dengan kemampuan tinggi (T) dengan peluang hanya 0,1%. Pada siswa yang kelompok prestasinya sangat rendah (SR) ini, kontribusi terbesar diberikan oleh faktor motivasi yang sangat rendah, minat belajar yang sangat rendah pula, serta latar belakang pendidikan guru yang tidak linear.

Tabel7 Peluangsiswa dengan latar belakang kurang baik

			Kelas1	Kelas2	Kelas3	Kelas4
Intersep			2,2619	1,8201	-	-
Lama_mqjr	(x_1)	(0;2]	0,4059	0,1715	-	0.2971
Tk_pend_gr	(x_2)	Lulus	0,0234	-	-	1.4056
Lb_pend_gr	(x_3)	Tidak	1,4025	0,8739	1.0139	-
Nyaman_gr	(x_4)	Sangat	-	0,2372	1.1297	-1.292
Byk_buku	(x_5)	[0;11)	0,0436	0,3515	-	-
Ekonm_ot	(x_6)	Sangat	0,3615	-	-	0.0645

Pend_ot	(x_7)	Sangat	0,7324	0,3992	-	-
Motivasi	(x_8)	Sangat	2,9732	2,1758	-	-
Persps_mat	(x_9)	Jelek	0,6965	0,2807	-	-
Minat_bljr	(x_{10})	Sangat	2,7912	-	-	0.4632
Intens_pr	(x_{12})	< 1 kali	0,79	0,1078	0.1307	-
Waktu_pr	(x_{13})	1-15	0,7835	0,0643	-	-
Prilaku_sis	(x_{14})	Sangat	-1,026	0,4436	0.2	0.3824
Reinforcmn	(x_{15})	Rendah	0,7157	0,1427	0.2593	-
Sarana_skl	(x_{16})	Sangat	-	-	0.0642	0.38
Waktu_lng	(x_{17})	Sangat	-	2,4149	-	4.6918
$\pi_k(x_i) = \frac{\exp(\beta_k x_i)}{\sum_{k=1}^4 \exp(\beta_k x_i)}$			0.9997	0.9972	0.0007	0.0010

Sebaliknya, siswa yang memiliki latar belakang yang baik sebagaimana terlihat pada Tabel 8, memiliki peluang yang besar untuk masuk ke dalam kelompok siswa yang prestasinya tinggi (T) dengan peluang sebesar 94,39% dan kecil peluangnya untuk masuk ke dalam kelompok siswa yang prestasinya sangat rendah (SR) yaitu sebesar 12,96%. Pada siswa yang kelompok prestasinya tinggi (T) ini, kontribusi terbesar diberikan oleh faktor latar belakang pendidikan guru yang linear, persepsi guru terhadap sekolah/tingkat kenyamanan guru di lingkungan sekolah, serta lama mengajar guru atau guru yang berpengalaman.

Tabel 8 Peluang siswa dengan latar belakang baik

			Kelas1	Kelas2	Kelas3	Kelas4
Intersep			22,619	18,201	-0.238	-3.844
Lama_mgjr	(x_1)	(14;20] tahun	-0.437	-0.392	-0.302	1.1315
Tk_pend_gr	(x_2)	Lulus S1	-0.657	-0.004	0.3257	0.3355
Lb_pend_gr	(x_3)	Linear	-0.942	-0.556	-0.361	1.8581
Nyaman_gr	(x_4)	Tinggi	-0.279	-0.371	-0.585	1.2344
Byk_buku	(x_5)	[200; >) buku	0.0225	-0.908	0.3895	0.4956
Ekonom_ot	(x_6)	Sangat Tinggi	-0.593	-0.08	0.062	0.6117

Pend_ot	(x_7)	Sangat Tinggi	-0.457	0.0405	0.2899	0.127
Motivasi	(x_8)	Sangat Tinggi	-0.212	-0.543	0.6744	0.08
Persps_mat	(x_9)	Baik	-0.725	-0.219	0.2329	0.711
Minat_bljr	(x_{10})	Sangat Tinggi	-0.698	0.3735	0.102	0.2222
Intens_pr	(x_{12})	Setiap hari	-0.294	-0.012	-0.011	0.3169
Waktu_pr	(x_{13})	> 90 menit	-0.119	0.0335	-0.172	0.2575
Prilaku_sis	(x_{14})	Baik	0.1024	-0.567	-0.279	0.7437
Reinforcmn	(x_{15})	Tinggi	-0.489	-0.14	-0.079	0.7077
Sarana_skl	(x_{16})	Sangat Banyak	0.1335	0.0607	-0.034	-0.16
Waktu_Ing	(x_{17})	Normal	1.477	-0.491	1.021	-2.007
$\pi_k(x_h) = \frac{\exp(\beta_k x_h)}{\sum_{k=1}^4 \exp(\beta_k x_h)}$			0.1296	0.1242	0.7382	0.9439

F. KESIMPULAN

Dari uraian di atas maka dapat disimpulkan ;

1. Analisis logistik kelas laten sebagai bagian dari analisis statistika khususnya analisis klastercocok untuk digunakan dalam mengelompokkan prestasi siswa dengan berbagai faktor atau latar-belakanyang mempengaruhinya.
2. Berdasarkan hasil analisis logistik kelas laten terhadap data hasil survey TIMSS tahun 2007, maka prestasi matematika siswa SLTP kelas 8 di Indonesia dapat digolongkan ke dalam empat kategori yaitu: *sangat rendah*, *rendah*, *sedang*, dan *tinggi* dengan proporsi masing-masing sebesar 39,16%, 32,42%, 21,46% dan 6,97%.
3. Latar belakang keempat kelompok siswa tersebut cukup baik dijelaskan oleh faktor-faktor seperti: berapa lama guru mengajar, program studi yang ditempuh guru, tingkat pendidikan guru, persepsi guru terhadap sekolah,

tingkat pendidikan orang tua, keadaan ekonomi orang tua, banyaknya buku pelajaran yang dimiliki, minat, motivasi, dan persepsi siswa terhadap matematika, perilaku siswa di sekolah, intensitas pemberian PR, banyaknya waktu untuk mengerjakan PR, sarana dan prasarana sekolah, serta pemanfaatan waktu luang.

4. Latar belakang siswa pada masing-masing kelompok tersebut berbeda-beda, siswa dengan latar belakang yang kurang baik, kecil sekali peluangnya untuk mencapai prestasi yang tinggi, dan sebaliknya siswa dengan latar belakang yang baik, besar sekali peluangnya untuk mencapai prestasi yang tinggi.

BIBLIOGRAFI

- Chung H. *et al.* Latent Class Logistic Regression: Application to marijuana use and attitudes among high school seniors. *Journal of The Royal Statistical Society*, 2006
- Hagenaars JA. *Latent Structure Models with direct effect between Indikator: Locals dependence models.* Sociological Methods and Research, 1998
- Lazarsfelds PF, Henry NW. *Latent Structur Analysis.* Boston: Houghton Mifflin. 1998
- Mullis IVS. *et al.* TIMSS 2007: *Assessment Frameworks from IEA's Trends in International Mathematics and Science Study at the fourth and Eight Grades.* MA, Boston, TIMSS InternationalStudyCenter: BostonCollege, 2005
- Nainggolan BMH. *Perbandingan Analisis Laten Kelas dengan Kriteria WHO untuk Penggerombolan Pasien Demam Dengue (DD) dan Demam Berdarah Dengue (DBD)*[tesis]. Bogor: Fakultas Matematika dan Ilmu Pengetahuan Alam, Intstitut Pertanian Bogor, 2009
- Santoso A. *Faktor-faktor yang mempengaruhi prestasi siswa: tinjauan berdasarkan data TIMSS 2007.* *Prosiding Seminar Mutu pendidikan dasar dan menengah hasil penelitian Puspendik.* Jakarta 28-29 Oktober 2009. Jakarta: Badan Penelitian dan Pengembangan Departemen Pendidikan Nasional, 2009
- Syah M. *Psikologi Pendidikan dengan Pendekatan Baru (Edisi Revisi).* Bandung: PT Remaja Rosda Karya, 2005

Pengelompokan Prestasi Matematika Siswa Indonesia Berdasarkan Hasil Survey Timss Menggunakan Analisis Logistik Kelas Laten

Vermunt JK, Magidson J. *Latent Class Cluster Analysis*. Hagnaars JA & McCutcheon AL (eds.) *Applied Latent Class Analysis*. Cambridge University Press, 2002