



**PENGEMBANGAN *WEB INTERFACE* UNTUK ANALISIS
ITEM RESPONSE THEORY DENGAN PENDEKATAN MODEL
CAMPURAN DAN PENERAPANNYA PADA BANK SOAL**

MGMP

TESIS

Oleh

Tika Clarinta Putri Utama

NIM 161820101004

**MAGISTER MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS JEMBER**

2018



**PENGEMBANGAN *WEB INTERFACE* UNTUK ANALISIS
ITEM RESPONSE THEORY DENGAN PENDEKATAN MODEL
CAMPURAN DAN PENERAPANNYA PADA BANK SOAL**

MGMP

TESIS

Oleh

Tika Clarinta Putri Utama

NIM 161820101004

**MAGISTER MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS JEMBER**

2018

PERSEMBAHAN

Dengan menyebut nama Allah Yang Maha Pengasih lagi Maha Penyayang, serta shalawat atas Nabi Muhammad SAW, tesis ini saya persembahkan untuk:

1. Ayahanda Suparto dan Ibunda Endang Nurhayati yang selalu mendoakan serta memberikan dukungan, semangat dan motivasi demi terselesaikannya tesis ini.
2. Kedua suami tercinta Vian Vidian Syah serta adik tersayang Tiara Miladiatul Kasana yang selalu memberikan solusi dalam setiap permasalahan.
3. Prof.Drs. I Made Tirta,M.Sc.,Ph.D selaku Dosen Pembimbing Utama dan Dr.Moh Fatekurrahman,S.Si.,Msi selaku Dosen Pembimbing Anggota yang senantiasa penuh kesabaran selama membimbing dalam menyelesaikan tugas akhir ini.
4. Dr. Alfian Futuhul Hadi, S.Si., M.Si dan Drs. Rusli Hidayat, M.Sc selaku dosen penguji yang senantiasa memberikan kritik dan saran demi kesempurnaan tugas akhir ini.
5. Seluruh guru SDN Tegaldlimo 1, SMPN 1 Purwoharjo dan SMA N 1 Purwoharjo yang telah memberikan banyak ilmu.
6. Seluruh dosen Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Jember yang membimbing penulis selama masa kuliah.
7. Almamater Jurusan Pendidikan Matematika Fakultas Keguruan dan Ilmu Pendidikan Universitas Muhammadiyah Malang.

MOTTO

"Barang siapa yang bersungguh sungguh, sesungguhnya kesungguhan tersebut untuk kebaikan dirinya sendiri"

(Qs. Al-Ankabut: 6)*)

"Lidahmu jangan kamu biarkan menyebut kekurangan orang lain, sebab kamu pun punya kekurangan dan orang lain pun punya lidah."

(Imam Syafii)**)

"Balas dendam terbaik adalah dengan memperbaiki dirimu."

(Ali Bin Abi Thalib)***)

*) Departemen Agama Republik Indonesia. 1998. Al-Qur'an dan Terjemahannya. Semarang: PT.Kumudasmoro Grafindo

***) <https://katakata.pro/kata-kata-mutiara/>

****) <https://www.kepogaul.com/inspirasi/motto-hidup-singkat-tapi-bermakna/>

PERNYATAAN

Saya yang bertanda tangan dibawah ini:

Nama : Tika Clarinta Putri Utama

NIM : 161820101004

menyatakan dengan sesungguhnya bahwa karya ilmiah yang berjudul “Pengembangan *WebInterface* untuk Analisis *Item Response Theory* dengan Pendekatan Model Campuran dan Penerapannya pada Bank Soal MGMP” adalah benar-benar hasil karya sendiri, kecuali kutipan yang sudah saya sebutkan sumbernya, belum pernah diajukan pada institusi manapun, dan bukan karya jiplakan. Saya bertanggung jawab atas keabsahan dan kebenaran isinya sesuai dengan sikap ilmiah yang harus dijunjung tinggi.

Demikian pernyataan ini saya buat dengan sebenarnya, tanpa ada tekanan dan paksaan dari pihak manapun serta bersedia mendapat sanksi akademik jika ternyata di kemudian hari pernyataan ini tidak benar.

Jember, May 2018,

Yang Menyatakan,

Tika Clarinta Putri Utama

NIM 161820101011

TESIS

Pengembangan *WebInterface* untuk Analisis *Item Response Theory* dengan Pendekatan Model Campuran dan Penerapannya pada Bank Soal MGMP

Oleh

Tika Clarinta Putri Utama

NIM 161820101004

Pembimbing

Dosen Pembimbing Utama : Prof.Drs. I Made Tirta,M.Sc.,Ph.D Dosen

Pembimbing Anggota : Dr. Mohamad Fatekurohman, S.Si., M.Si

PERSETUJUAN PEMBIMBING

Tesis yang berjudul “ Pengembangan *WebInterface* untuk Analisis *Item Response Theory* dengan Pendekatan Model Campuran dan Penerapannya pada Bank Soal MGMP” telah disetujui pada:

hari, tanggal :

tempat : Fakultas Matematika dan Ilmu Pengetahuan Alam

Tim Penguji:

Dosen Pembimbing Utama,

Dosen Pembimbing Anggota ,

Prof.Drs. I Made Tirta,M.Sc.,Ph.D
NIP 195912201985031002

Dr.Moh Fatekurrahman,S.Si.,Msi
NIP 196906061998031001

Dosen Penguji I,

Dosen Penguji II,

Dr. Alfian Futuhul Hadi, S.Si., M.Si
NIP 197407192000121001

Drs. Rusli Hidayat, M.Sc
NIP 196610121993031001

Mengesahkan,
Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Jember

Drs. Sujito, Ph.D.

NIP. 196102041987111001

RINGKASAN

Pengembangan *WebInterface* untuk Analisis *Item Response Theory* dengan Pendekatan Model Campuran dan Penerapannya pada Bank Soal MGMP; Tika Clarinta Putri Utama, 161820101004; 2018: 42 halaman; Program Studi Magister Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Jember.

Dunia pendidikan dewasa ini mengalami kemajuan yang sangatlah pesat. Guru harus memiliki kompetensi yang memadai dalam setiap proses kegiatan belajar mengajar. Guru dituntut untuk mampu melaksanakan proses belajar mengajar yang sesuai dan bervariasi dari setiap materi ajar. Dengan memiliki kompetensi yang cukup maka seorang guru mampu mengajar dan mengevaluasi dengan baik. Hal tersebut tidaklah mudah, untuk itu guru-guru mempunyai wadah yang disebut MGMP (Musyawarah Guru Mata Pelajaran) guna membantu mengembangkan kompetensi guru yang ada. Fungsi MGMP disini salah satunya adalah untuk menampung aspirasi guru dan kesulitan guru dalam mempersiapkan. Karena pentingnya pengembangan uji validasi yang selama ini masih menggunakan uji tes tradisional. Agar mendapatkan hasil bank soal yang valid, maka perlunya mengadopsi instrumen pengujian bank soal yang baru, dan menyajikannya dengan lebih mudah.

Penyusunan tes dapat dilakukan dimulai dari perancangan, perakitan atau pembuatan dan analisis butir tes dengan menggunakan berbagai pendekatan, baik menggunakan teori tes klasik atau disebut juga *Classical Test Theory (CTT)*, teori tes modern disebut juga *Item Response Theory (IRT)*, sedang teori tes *Multidimensional Item Response Theory (MIRT)*. Pengembangan tes dengan pendekatan teori respon butir berkembang sangat pesat dan memiliki keunggulan.

Kemajuan teknologi informasi sudah banyak dimanfaatkan dalam mendukung kegiatan pembelajaran dan perlunya perancangan yang sangat baik dalam penyusunan perangkat maupun dalam pemanfaatnya pada bank soal *MGMP*. Oleh sebab itu dibutuhkan alat bantu untuk mempermudah guru

mengevaluasi dalam setiap pembelajaran. Terbatasnya ketersediaan piranti lunak untuk menganalisis soal memanfaatkan teori respon butir *IRT*. Memiliki keunggulan. “Dalam melakukan simulasi *R* melalui *R-shiny*, bagian ini merupakan otak dari program yang bertugas melakukan simulasi, berbagai analisis data sesuai pilihan pengguna dan selanjutnya mengirim hasilnya ke bagian output, bagian ini didukung oleh berbagai prosedur dan analisis data yang pada umumnya telah tersedia pada berbagai paket *R* bekerja di server (belakang layar) sedang penggunaan disuguhkan menu yang ramah dan interaktif berbasis web, sehingga pengguna tidak disyaratkan menguasai program *R*. Oleh karena itu *R-shiny* sangat baik digunakan sebagai pelengkap pembelajaran statistika dimana pengguna bisa lebih fokus pada konsep statistika tanpa terkendala kemampuan penguasaan *software* statistika” (Tirta, 2014). Akan tetapi, cara penggunaannya relatif sulit karena berbasis pemrograman atau skrip. Diperlukan program tambahan yang dapat membuat analisis respon butir mudah dilakukan dengan *R*, salah satu yang bisa dilakukan adalah membuat interface berbasis *R-shiny*. Cara ini sudah dilakukan oleh beberapa peneliti untuk kebutuhan yang lain dengan memanfaatkan *R-shiny*. (Tirta, 2015. Tirta et al 2017, Tirta & Dian 2018). Program *R* ini adalah dasar dari program *software* pengolah data *web interface* yang disajikan dalam penelitian ini untuk menganalisis butir soal MGMP.

Differential Item Functioning atau (*DIF*) dapat dianggap tidak bias, jika item tersebut memang mengukur bagian dari kemampuan yang sesuai dengan tujuan pengukuran, hanya saja salah satu kelompok, meskipun memiliki abilitas yang sama, cenderung memiliki kelemahan dalam bagian tersebut. Dampak merupakan pengaruh dari perbedaan kelompok dalam atribut yang diukur pada hasil tes. Ketika perbedaan antar kelompok terjadi mungkin karena adanya perbedaan kuantitas atribut yang diukur dari kedua kelompok, maka tes atau item tersebut dianggap tidak mengalami *DIF*. Perbedaan skor dua subjek yang berasal dari dua kelompok yang berbeda dapat terjadi karena memang keduanya memiliki kuantitas atribut yang berbeda atau karena item mengalami *DIF*. Item yang baik seharusnya berlaku adalah bagi setiap peserta tes, tidak boleh menguntungkan

suatu kelompok tertentu seperti menguntungkan kelompok perempuan, sehingga merugikan kaum laki – laki. Faktor – faktor yang mempengaruhi *DIF* dibedakan atas faktor internal dan faktor eksternal.

Penelitian dilaksanakan di 3 Sekolah dengan menyebarkan 50 soal yang diambil secara acak sedangkan pengembangan web dilakukan di Laboratorium Statistika Universitas Jember Fakultas FMIPA dengan mengikuti proses tahap penelitian dan uji analisis validasi bank soal MGMP. Soal yang dianalisis adalah bagian dari bank soal yang disusun oleh Tim MGMP termasuk penulis sendiri, meliputi lingkup, jenis soal dan banyaknya soal. Sampel soal dalam penelitian ini mengambil sample soal dari 150 soal pilihan ganda, yang dimiliki bank soal MGMP Matematika di Kabupaten Banyuwangi. Soal yang digunakan dalam uji penelitian ini berjumlah 50 soal pilihan ganda. Sebelum melakukan uji tes peneliti memberikan kisi – kisi soal terlebih dahulu. Jumlah soal yang disediakan pada tiap lembaga sekolah adalah 50 pilihan ganda dengan waktu 120 menit, dalam setiap soal terdiri dari 10 soal berbobot mudah, 30 soal berbobot sedang dan 10 soal berbobot sulit, setiap soal memiliki jawaban pengecoh sebanyak 5 item.

Subjek dalam penelitian ini berasal dari Siswa Menengah Atas di Kabupaten Banyuwangidan dilaksanakan di 3 sekolah yaitu SMA PGRI Tegaldlimo, SMA NEGRI 1 Purwoharjo, dan SMA NU Genteng. Adapun data yang diperoleh dalam peneliti ini mewakili beberapa cluster atau golongan kelompok dalam populasi penelitian ini berupa:

- a) Golongan menurut status lembaga negeri dan swasta kelompok sekolah
- b) Golongan menurut tingkat sejajar XII IPA dan,
- c) Golongan menurut lokasi atau wilayah yang berada di Kabupaten Banyuwangi

Penyusunan *Web Interface*

Web interface dikembangkan dengan menggunakan *R-shiny* dalam format pirati lunak (bukan modul) yang didukung oleh dua file utama yaitu *ui.r* (untuk memprogramkan menu interaktif dengan pengguna) dan *server.r* untuk mengirim

semua perintah yang diperlukan ke R (dalam bentuk server). Kemampuan / fitur menu yang dicakup meliputi :

- a) Pengolahan skor kasar ke skor biner
- b) Analisis IRT (satu-3 PL)
- c) Analisis Multi Dimensional
- d) Aplikasi HGLM dalam IRT

Maka dapat disimpulkan bahwa dari hasil penelitian ini terdapat duahal yang bermanfaat yaitu:

1. *Web Interface* memudahkan dengan tampilan fitur – fitur menu yang mamprediksi tiap item soal yang memiliki kekurangan atau perlunya evaluasi pada tiap soal yang dinyatakan kurang baik. Serta cara kerja yang mempermudah guru memasukkan banyak data ini menunjukan web interface lebih efisien dibandingkan dengan yang berbasis MS. Excel.
2. Dari hasil analisis uji coba bank soal MGMP di Banyuwangi, terdapat 19 soal yang dinyatakan kurang baik atau adanya revisi pada soal nomor nomor 2, 5, 8, 9, 10, 13, 17, 19, 21, 23, 24, 24, 29, 34, 37, 38, 40, 42 dan 48. Perlunnya evaluasi pada tiap soal yang mengalami tingkat kesulitan yang tinggi dan tingkat kesulitan yang rendah. Dengan efek tetap Nilai bahasa Indonesia dan Efek Gender, Sekolah (sekolah dianggap acak mewakili sekolah di Banyuwangi) Hasil kontribusi Gender dan Bhs Indonesia tidak cukup signifikan untuk menjawab pertanyaan dengan benar. Ini juga berarti tidak bias (*DIF*). Jenis kelamin meskipun tidak signifikan, jika semuanya dipertahankan dalam model maka diperoleh hasil yang baik. Karena dalam keseluruhan soal tidak mengandung *DIF* pada tiap butir soal yang di buat oleh MGMP.

KATA PENGANTAR

Puji dan syukur penulis panjatkan kehadirat Allah SWT yang telah memberikan rahmat, taufik serta hidayah-Nya sehingga tesis yang berjudul “Pengembangan *WebInterface* untuk Analisis *Item Response Theory* dengan Pendekatan Model Campuran dan Penerapannya pada Bank Soal MGMP” dapat terselesaikan. Tugas akhir ini disusun untuk memenuhi syarat dalam menyelesaikan pendidikan strata dua (S2) pada Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Jember.

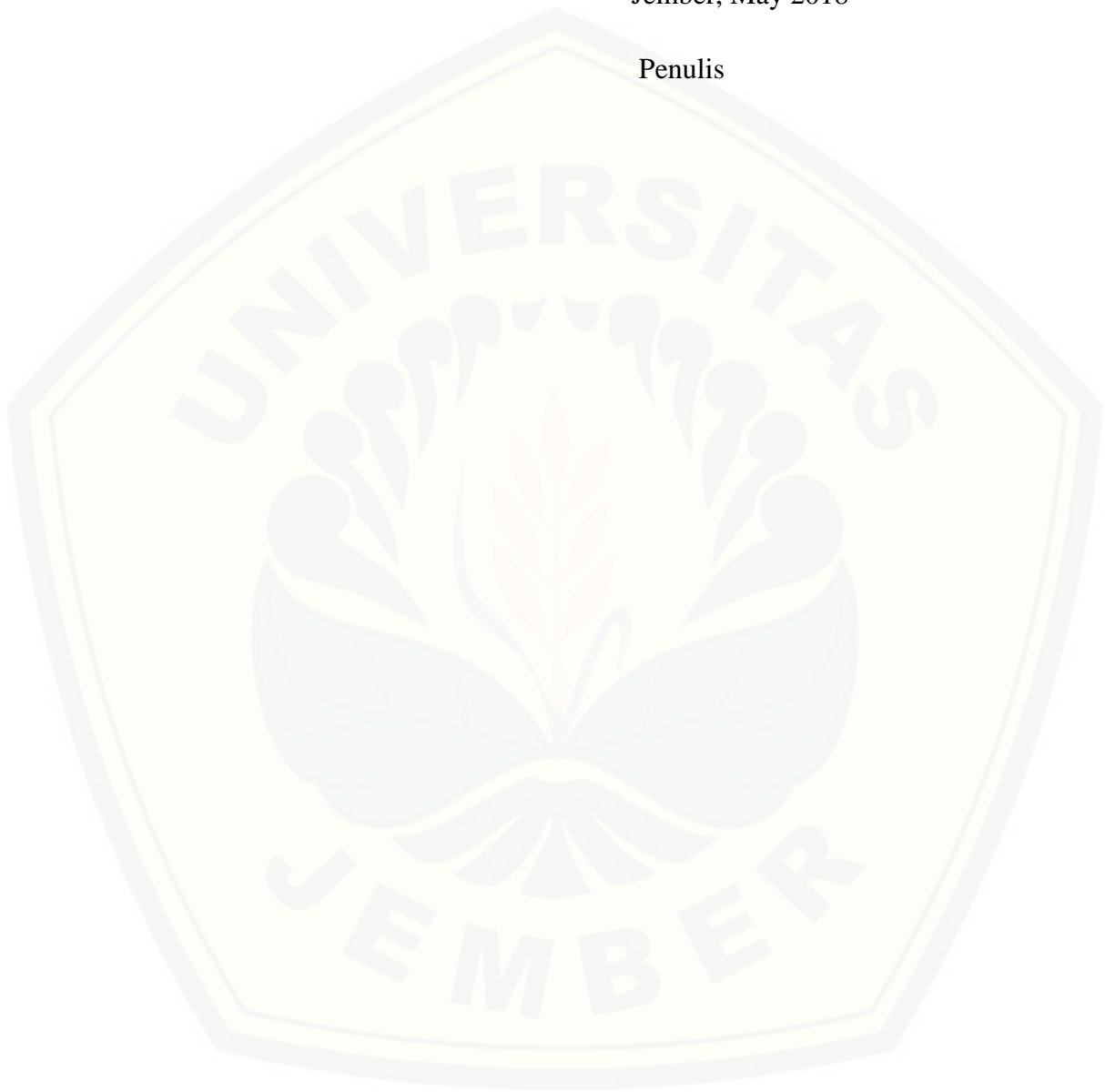
Penyusunan tesis ini tidak terlepas dari bantuan berbagai pihak, baik secara langsung maupun tidak langsung. Oleh karena itu, penulis menyampaikan terima kasih kepada:

1. Prof.Drs. I Made Tirta, M.Sc., Ph.D selaku Dosen Pembimbing Utama dan Dr. Mohamad Fatekurohman, S.Si., M.Si. selaku Dosen Pembimbing Anggota yang telah meluangkan waktu, pikiran, dan perhatian dalam penulisan tesis ini;
2. Dr. Alfian Futuhul Hadi, S.Si, M.Si. selaku Dosen Penguji I dan Drs. Rusli Hidayat, M.Sc selaku dosen penguji II yang telah memberikan kritik dan saran demi kesempurnaan tesis ini;
3. Kusbudiono, S.Si., M.Si. selaku ketua Jurusan Matematika Fakultas MIPA Universitas Jember;
4. Seluruh staf pengajar Jurusan Matematika Fakultas MIPA Universitas Jember yang telah memberikan ilmu serta bimbingannya sehingga penulis dapat menyelesaikan tesis ini;
5. Keluarga Besar Mahasiswa Magister Matematika 2016 yang telah memberikan semangat dan bantuan dalam proses penulisan tesis ini;
6. Semua pihak yang tidak dapat disebutkan satu-persatu.

Penulis juga menerima segala kritik dan saran dari semua pihak demi kesempurnaan tesis ini. Akhirnya penulis berharap, semoga tesis ini dapat bermanfaat.

Jember, May 2018

Penulis



DAFTAR ISI

	Halaman
HALAMAN JUDUL	i
HALAMAN PERSEMBAHAN	ii
HALAMAN MOTTO	iii
HALAMAN PERNYATAAN.....	iv
HALAMAN PEMBIMBINGAN.....	v
HALAMAN PENGESAHAN.....	vi
HALAMAN PEMBIMBINGAN.....	v
RINGKASAN	vii
KATA PENGANTAR.....	xi
DAFTAR ISI.....	xiii
DAFTAR GAMBAR.....	xv
DAFTAR TABEL	xvi
DAFTAR SINGKATAN.....	xvii
DAFTAR LAMPIRAN.....	xviii
BAB 1. PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	4
1.3 Tujuan Penelitian	4
1.4 Manfaat Penelitian	4
BAB 2. TINJAUAN PUSTAKA.....	5
2.1 Kualitas Pengujian Bank Soal.....	5
2.2 Teori Tes Kalasik.....	6
2.2.1 Tingkat Kesulitan Item.....	6
2.2.2 Daya Beda Item.....	7
2.3 Teori Tes Modern.....	8
2.4 Item Response Theory(Unidimensi)	9
2.5 Multidimensional Item Response Theory atau MIRT	11
2.6 Pengertian DIF (Differential Item Functioning)	

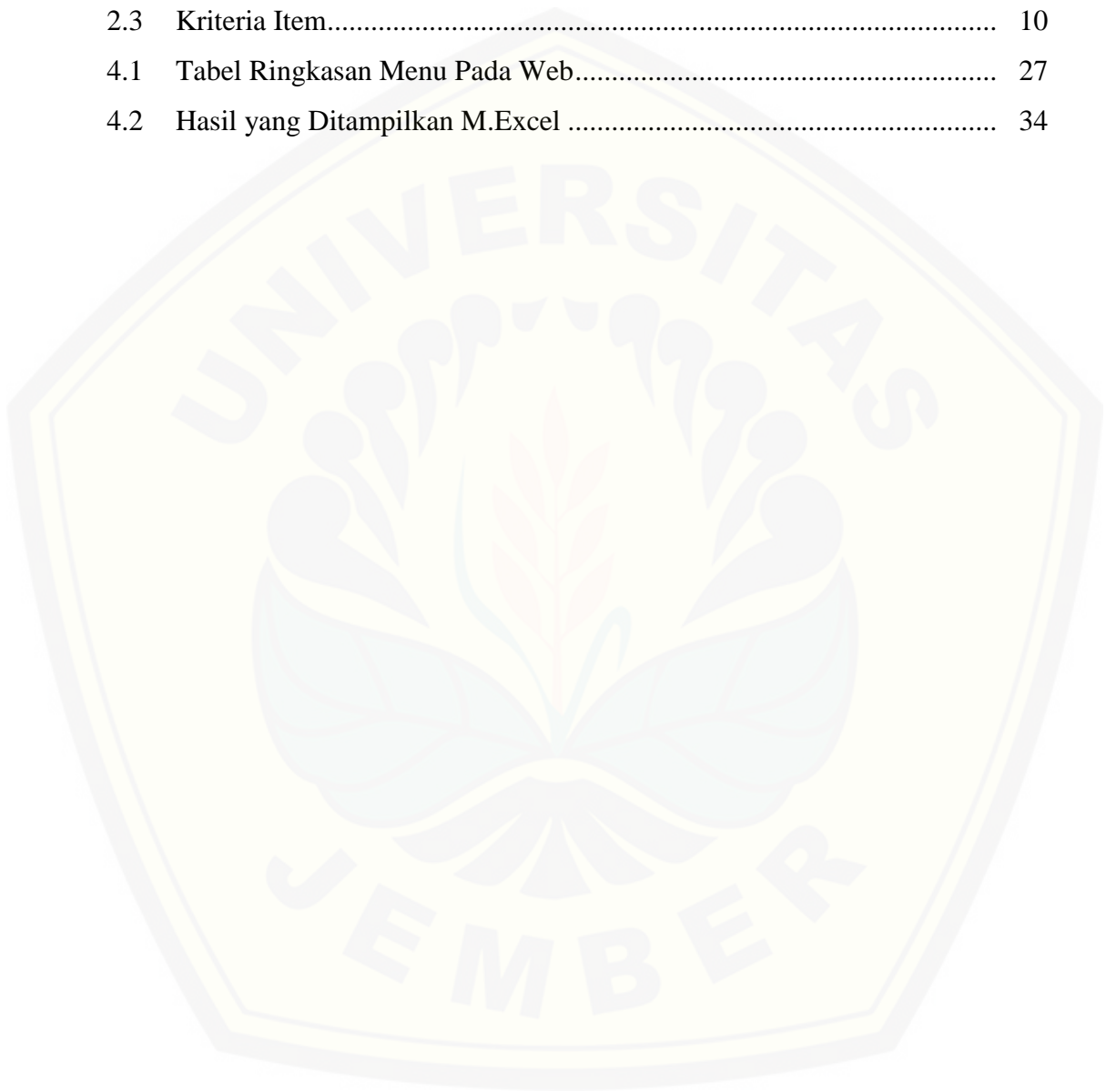
Dan Pendekatan <i>HGLM</i>	14
2.7 Struktur Pemrograman Simulasi R-Shiny	20
BAB 3. METODE PENELITIAN	22
3.1 Data Penelitian	22
3.2 Tahap Penelitian	22
3.2.1 Penyusunan Web Interface.....	21
3.2.2. Penerapan Data Bank Soal	23
3.3 Observasi Awal	24
BAB 4. HASIL DAN PEMBAHASAN	25
4.1 Subjek Penelitian	25
4.1.1 Sampel Siswa	26
4.1.2 Sampel Soal.....	26
4.1.3 Tampilan <i>Web Interface</i>	26
4.2 Analisis Data	29
4.2.1 Eksplorasi Data	29
4.2.2 Unidimensional	29
4.2.3 Multidimensional	31
4.3 Pembahasan	32
4.3.1 Kinerja Performance Program Web	32
4.3.2 <i>Item Response Theory</i>	35
4.3.3 <i>Multidimensional Item Response Theory</i>	36
4.3.4 Pendekatan <i>HGLM</i>	37
BAB 5. KRITIK DAN SARAN	38
5.1 Kesimpulan	38
5.2 Saran	38
DAFTAR PUSTAKA	39

DAFTAR GAMBAR

	Halaman
2.7 Struktur Komponen R – Shiny	21
3.1 Diagram alir tahap penelitian pelaksanaan validasi bank soal MGMP Banyuwangi	24
3.2 Diagram alir tahap pengujian dalam bentuk <i>Web Interface</i>	24
3.3 Tahapan Progres dan Progress Kegiatan Penelitian	25
4.1 Data Asli	29
4.2 Data Score	29
4.3 Item Characteristic Curve	31
4.4 Nilai Tiap Item	33
4.5 Tingkat Kesulitan Item.....	34

DAFTAR TABEL

	Halaman
2.1 Skala Tingkat Kesukaran Butir.....	7
2.2 Interval Daya Beda	8
2.3 Kriteria Item.....	10
4.1 Tabel Ringkasan Menu Pada Web.....	27
4.2 Hasil yang Ditampilkan M.Excel	34

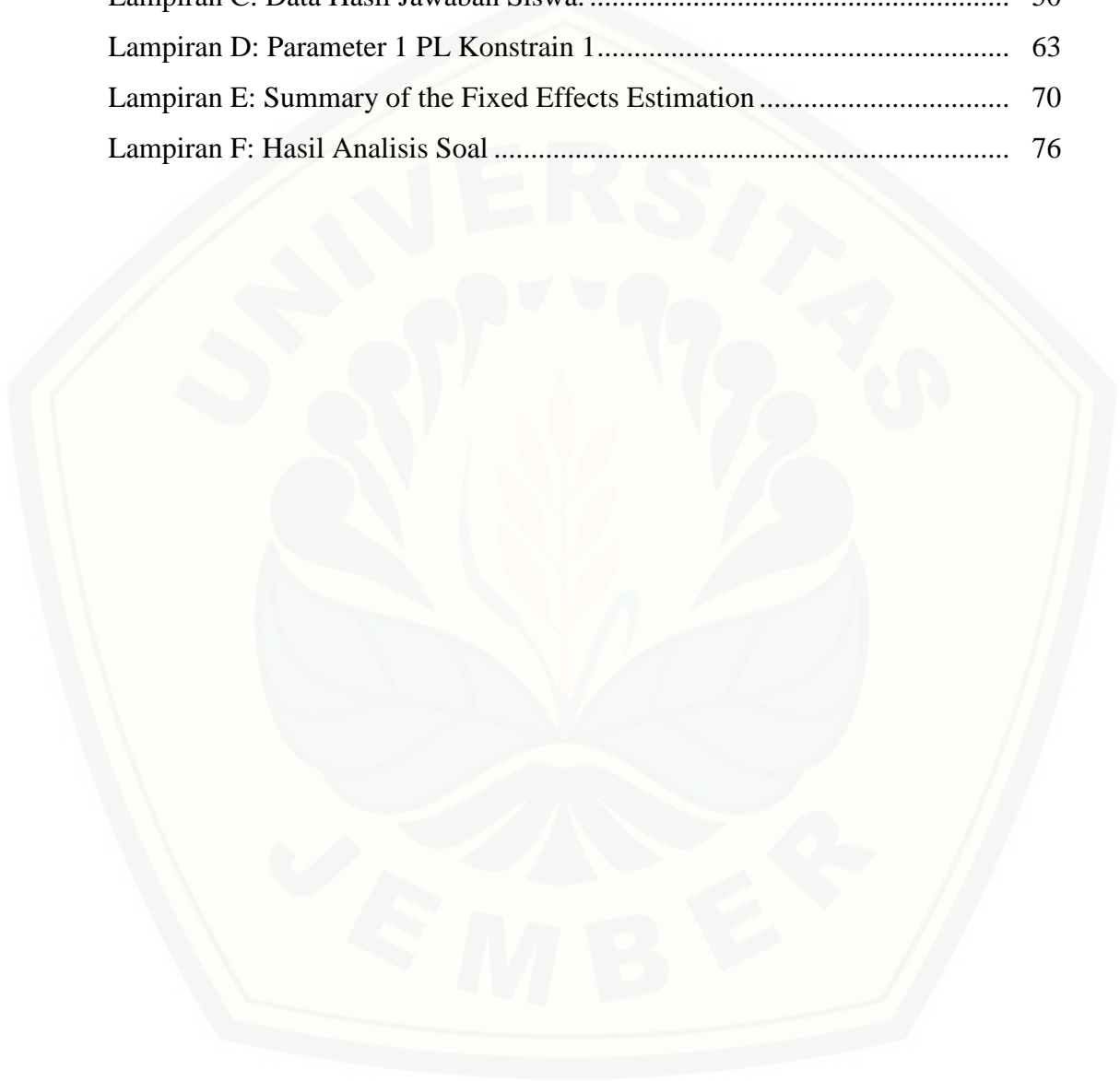


DAFTAR SINGKATAN

MGMP	: Musyawarah Guru Mata Pelajaran
IRT	: <i>Item Response Theory</i>
CTT	: <i>Classical Test Theory</i>
ICC	: <i>Item Characteristic Curve</i>
MIRT	: <i>Multidimensional Item Response Theory</i>
TPA	: Tes Potensi Akademik
GLMM	: <i>Generalized Linear Mixed Model</i>
GAISE	: <i>Guidelines for Assessment and Instruction in Statistics Education</i>
HGLM	: <i>Hierarchical Generalized Linear</i>

DAFTAR LAMPIRAN

	Halaman
Lampiran A: Data Siswa	42
Lampiran B: Ringkasan Data pada <i>Interface</i>	47
Lampiran C: Data Hasil Jawaban Siswa	50
Lampiran D: Parameter 1 PL Konstrain 1	63
Lampiran E: Summary of the Fixed Effects Estimation	70
Lampiran F: Hasil Analisis Soal	76



BAB 1. PENDAHULUAN

1.1 Latar Belakang

Perkembangan dunia pendidikan dewasa ini mengalami kemajuan yang sangatlah pesat, adanya era teknologi yang semakin modern, sehingga guru harus memiliki kompetensi yang memadai dalam setiap proses kegiatan belajar mengajar. Oleh sebab itu guru dituntut untuk mengembangkan pada dirinya sehingga mampu untuk melaksanakan proses belajar mengajar yang berbeda dan bervariasi dari setiap materi ajar, dengan memiliki kompetensi yang cukup maka seorang guru mampu mengajar dan mengevaluasi dengan baik. Hal tersebut tidaklah mudah, maka guru - guru mempunyai wadah yang disebut MGMP (Musyawarah Guru Mata Pelajaran) guna membantu mengembangkan kompetensi guru yang ada di Kabupaten Banyuwangi. Fungsi MGMP disini untuk menampung aspirasi guru dan kesulitan guru menghadapi kemajuan perkembangan tes yang semakin modern. Karena pentingnya pengembangan uji validasi yang selama ini penafsirannya masih menggunakan tes tradisional atau klasik, dengan menggunakan kemampuan para ahli dan secara logis saja tidak cukup, untuk mendapatkan hasil bank soal yang valid, maka perlunya mengadopsi instrumen pengujian bank soal yang baru, dan menyajikannya dengan lebih mudah.

Terkait pembelajaran statistika, sebelum perkembangannya komputer, ilustrasi pengajaran komputer yang pesat, visualisasi dan ilustrasi bisa digantikan dengan simulasi komputer yang banyak tersedia secara online (Blejec,2003). Secara khusus American Statistic Association dalam laporannya, yang terkenal dengan GAISE “*Guidelines for Assessment and Instruction in Statistics Education*” pada tahun 2000 (Kahle,2014). Merekomendasikan 6 hal terkait pembelajaran statistika, tiga diantaranya bahwa pembelajaran statistika harus memenuhi (i) *stress conceptual understanding rather than more knowledge of procedure*; (ii) *foster active learning in the classroom*; (iii) *use technology for developing conceptual understanding and analyzing data*. Laporan ini mengindikasikan perlunya pemanfaatan teknologi dalam pembelajaran statistika

yang dapat mendorong siswa menjadi pembelajar aktif, membantu mereka lebih memahami konsep dan tidak sekedar mendapat pengetahuan menghitung. Visualisasi menggunakan *R* yang pada dasarnya merupakan software pengolah data, memiliki kelebihan dalam simulasi dan visualisasi bisa dilakukan pada semua aspek analisis data (simulasi data, estimasi, dan pemeriksaan *goodnes of fit*) dalam Tirta, (Sanchez, 2013).

Untuk mengetahui kemampuan seseorang di bidang tertentu, tes merupakan salah satu instrumen yang sering digunakan. Tes dapat digunakan setelah melalui proses penyusunan. Penyusunan tes dapat dilakukan dimulai dari perancangan, pembuatan dan analisis butir tes dengan menggunakan berbagai pendekatan, baik menggunakan teori tes klasik atau teori respon butir. Pendekatan dengan pendekatan teori respon butir memiliki keunggulan. Model teori respon butir menghasilkan parameter butir yang independen dari peserta tes dan parameter peserta tes yang bersifat independen dari sekumpulan tes yang diujikan (Xitao, 1998), karena model logistik didasarkan pada fungsi respon logistik, dan efek acak diasumsikan terdistribusi normal model, yang terkait erat ini sering disebut sebagai model logistik normal, terutama dalam model laten, demikian pula model normal. Model konfirmasi dapat diestimasi dengan metode kuadratur, fungsi untuk memperkirakan model 2 dan 3 parameter diimplementasikan serta model respons bergradasi untuk data *polytomous*, dan model logistik multidimensional linier. *IRT-Shiny* menyediakan aplikasi mengkilap interaktif untuk analisis *IRT*. Paket *mcIRT* menyediakan fungsi untuk memperkirakan Model Respon Nominal dan Model Logit. Keduanya adalah model untuk memeriksa berbagai pilihan item dan format respons *polytomous* lainnya.

Dalam hal ini juga didukung penggunaan model campuran atau sering disebut *Mixed Models*. Model linier umum menunjukkan kelas model regresi efek tetap untuk kontinu dan dikotomis. (McCullagh dan Nelder) menjelaskan hal ini dengan sangat rinci dan menunjukkan bahwa istilah *GLM* (Nelder dan Wedderburn) yang menjelaskan bagaimana kumpulan teknik statistik yang tampaknya berbeda dapat disatukan. *Generalized Linear Model* umum meliputi regresi linier, regresi logistik dan regresi poisson.

Kemajuan teknologi informasi sudah banyak dimanfaatkan dalam mendukung kegiatan pembelajaran dan perlunya perancangan yang sangat baik dalam penyusunan perangkat maupun dalam pemanfaatannya pada bank soal *MGMP*. Oleh sebab itu dibutuhkan alat bantu untuk mempermudah guru dalam mengevaluasi dalam setiap pembelajaran. “Dalam melakukan simulasi R melalui *R-shiny* bekerja di server (belakang layar) sedang penggunaan disuguhkan menu yang ramah dan interaktif berbasis web, sehingga pengguna tidak disyaratkan menguasai program R. Oleh karena itu *R-shiny* sangat baik digunakan sebagai pelengkap pembelajaran statistika dimana pengguna bisa lebih fokus pada konsep statistika tanpa terkendala kemampuan penguasaan *software statistika*” (Tirta, 2014). Program R ini adalah dasar dari program *software* pengolah data yang disajikan dalam penelitian ini.

Dari beberapa pemaparan diatas perlu adanya perkembangan dan pembaharuan dalam pengujian bank soal yang di buat oleh MGMP di Banyuwangi. Mengingat pentingnya tes pengujian bank soal maka dibutuhkan pembaharuan dengan menggunakan *Item Response Theory* serta menyajikannya berupa *web interface*.

1.2 Rumusan Masalah

Dari latar belakang masalah, diperoleh rumusan masalah sebagai berikut :

- a. Bagaimana bentuk penyajian analisis *Item Response Theory* dengan pendekatan model campuran dalam *web interface*?
- b. Bagaimana hasil analisis yang dilakukan dengan menggunakan *Item Response Theory* dengan pendekatan model campuran pada pengujian bank soal MGMP?

1.3 Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah :

- a. Membuat penyajian analisis *Item Response Theory* dengan pendekatan model campuran dalam *web interface*.
- b. Menyajikan analisis *Item Response Theory* dengan pendekatan model campuran pada penerapan bank soal MGMP berbasis *web interface*.

1.4 Manfaat Penelitian

Dengan di adakan penelitian diharapkan akan dapat memberikan manfaat sebagai bahan rujukan dan literatur. Melalui penelitian ini diharapkan dapat membantu gurumenganalisis perkembangan siswa dengan penyajian *web interface*penerapannya pada bank soal MGMP di Banyuwangi.

Melalui tes uji dilapangan, diharapkan mampu menyajikan analisis yang dikembangkan dalam *web interface*.Penelitian ini akan mendapatkan pemahaman lebih tentang proses analisis *Item Respons Theory* dengan pendekatan model campuran pada penerapan bank soal dan juga cara pemakaian *web interface* yang mampu mempermudah guru memahami pengukuran tes dari tradisional *IRT* ke moderen dengan lebih efisien. Hal ini memotivasi peneliti dalam mengembangkan *web interface*yang dapat membantu guru – guru MGMP di Kabupaten Banyuwangi.

BAB 2. TINJAUAN PUSTAKA

Bab ini akan membahas dasar – dasar teori tentang pengukuran tes dan metode yang akan digunakan untuk menganalisis perlunya dan pentingnya bentuk penyajian analisis *IRT* dengan pendekatan model campuran dan penerapannya pada bank soal MGMP.

2.1 Kualitas Pengujian Bank Soal

Untuk Pengujian soal yang berbentuk tes, baik untuk pilihan maupun uraian, kualitas instrumen juga dilihat dari setiap butir soalnya. Dikatakan demikian, sebab secara keseluruhan mungkin bisa dinyatakan valid dan reliabel, namun belum tentu setiap butirnya juga baik. Untuk mengetahui kualitas butir tes, maka dibutuhkan program *web interface* yang dapat membantu proses validasi bank soal MGMP. Untuk itu kami para guru tergabung dalam *MGMP* (Musyawarah Guru Mata Pelajaran). *MGMP*, merupakan suatu organisasi guru yang dibentuk untuk menjadi forum komunikasi yang bertujuan untuk memecahkan masalah yang dihadapi guru dalam pelaksanaan tugasnya sehari-hari di lapangan. Peneliti berusaha agar semua data numerik hasil pengukuran yang dilakukan benar – benar merupakan data yang ingin diukur. Untuk itu pengujian bank soal yang digunakan harus memiliki validitas yang baik. Selain itu pengujian bank soal harus memiliki indeks reliabel. Tes sering digunakan dalam dunia pendidikan dan psikologi untuk mendapatkan data yang ingin diketahui dari individu atau kelompok tertentu, seperti seleksi pengambilan individu yang memiliki kriteria yang diinginkan. Bisa juga sebagai profile seseorang untuk mengetahui dibagian mana individu mampu mengetahui potensi pada dirinya dalam sebuah pekerjaan. Sebagai bahan pertimbangan pengukuran tes pengukuran sangatlah penting untuk terus dikembangkan untuk melihat kemajuan atau potensi seseorang dalam bidang tertentu. Untuk itu penyajian kualitas haruslah valid, objektif dan efisien.

Pada uji coba pengujian bank soal maka terdapat pula ragam program pengukuran.. Program pengukuran klasik berkembang dari (*IRT*) yang sudah banyak dikenal. Program sederhana dan mudah untuk menerapkan serta mengaplikasikannya namun masih banyak terdapat kelemahan. Untuk itu program pengukuran validasi diciptakan guna menutup kelemahan program klasik tersebut.

2.2 Teori Tes Klasik

Tes adalah suatu teknik pengukuran dirancang sebagai prosedur yang sistematis untuk mempelajari tingkah laku atau sekelompok individu(Cronbach,1970). Menurut Sudijono (2011), Dalam dunia pendidikan meliputi pengukuran kemampuan peserta tes dan pengukuran karakteristik alat ukur yang digunakan. Ada dua teori pengukuran yang sampai saat ini masih dikembangkan guna mempermudah pengukuran pada pendidikan,yaitu teori tes klasik dan teori tes modern. Teori klasik disebut juga (*CTT*), sedang teori tes modern disebut juga (*IRT*).

2.2.1 Tingkat Kesulitan Item

Dalam teori tes klasik yang sering digunakan dalam pengukuran tes ini, haruslah memenuhi asumsi – asumsi yang berlaku. Sebab teori ini tersusun berdasarkan asumsi – asumsi dari Allen &Yen (1979). Untuk mengetahui karakteristik taraf kesukaran butir, selanjutnya dilambangangkan dengan p , sering juga dituliskan sebagai *proportion of correct (p)* didefinisikan sebagai perbandingan antara jumlah jawaban benar dengan jumlah peserta tes. Atau dapat dituliskan dalam bentuk rumus matematik:

$$p = \frac{\sum B}{N} \quad (2.1)$$

p = indeks kesukaran soal

$\sum B$ = jumlah jawaban benar

N = jumlah peserta tes

Nilai batasan tingkat kesukaran soal terletak antara 0 dan 1, karena merupakan perbandingan antara jumlah jawaban benar dengan jumlah soal, dapat dituliskan dalam bentuk matematika berikut ini : $0 \leq p \leq 1$, bilangan

menunjukkan sukar dan mudahnya suatu soal disebut indeks kesukaran (Arikunto, 1999). Batasan – batasan tingkat kesukaran ini dapat dilihat dalam Tabel 2.1 berikut ini (Hayat, 1997):

Tabel 2.1 Skala Tingkat Kesukaran Butir

Indek Kesukaran (p)	Kategori Soal
$p > 0,70$	Mudah
$0,30 \leq p \leq 0,70$	Sedang
$p < 0,30$	Sulit

2.2.2 Daya Beda Item

Karakteristik daya beda (D = daya pembeda) butir soal. Didefinisikan sebagai kemampuan butir untuk membedakan antara kelompok yang pintar dan kelompok yang kurang mampu. Pengertian tes (bank soal) dalam membedakan antara mereka yang tinggi kemampuannya dan mereka yang rendah kemampuannya dalam hal – hal yang diukur oleh tes yang bersangkutan (Suryabrata, 1984). Jika dalam suatu kelompok peserta tes dibagi pada dua kelompok. Kelompok X dan Y maka daya beda soal adalah kemampuan soal itu membedakan antara kelompok X dan Y. Berikut formula yang dapat digunakan untuk menghitung indek daya pembeda (Allen & Yen. 1979), adalah :

$$r_{bis} = \frac{X_i - X}{S_x} \sqrt{\frac{x_i}{q_i}} \quad (2.2)$$

dengan r_{bis} = koefisien biserial

X_i = rata – rata nilai pada butir ke- i

X = rata – rata nilai total

S_x = standar deviasi nilai total

x_i = probabilitas jawaban benar

q_i = probabilitas jawaban salah ($1 - x_i$)

Rentang kolerasi skor butir dengan sekor peserta, yang baik adalah antara 0,40 – 0,70, sedangkan yang memiliki $< 0,2$ harus dibuang, dan memiliki daya

pembeda $> 0,70$ disebut sangat baik (Sudijono, 2001). Nilai negatif berarti daya bedanya tidak berfungsi. Lebih lengkap di sajikan pada Tabel 2.2.

Tabel: 2.2 Interfal Daya Beda Butir (D)

Interval (a)	Klasifikasi	Interprestasi
$D \leq 0,20$	Jelek	Daya Pembeda Jelek
$0,20 \leq D \leq 0,40$	Memuaskan	Memiliki Daya Pembeda Yang Cukup
$0,41 \leq D \leq 0,70$	Baik	Memiliki Daya Pembeda Yang Baik
$0,71 \leq D \leq 1,00$	Sangat Baik	Memiliki Daya Pembeda Yang Sangat Baik

Pada dasarnya batasan nilai a pada tabel 2.2 menunjukkan batasan yang cukup baik setelah dilakukan perbaikan pada batas – batas intervalnya. Kedua batas kiri dan batas kanan dari setiap kelas interval tidak saling terkait satu dengan yang lain.

2.3 Teori Tes Modern

Untuk mengatasi kelemahan – kelemahan yang ada pada teori klasik, para ahli pengukuran berusaha untuk mencari alternatif. Teori tes modern atau biasa disebut dengan teori butir (*IRT*) dikembangkan oleh para ahli pengukuran bidang pendidikan dan psikologi sebagai upayameminimalkan kekurangan – kekurangan yang ada dalam teori tes klasik. Perhitungan dalam analisis butir berdasarkan teori ini dapat dilakukan dengan menggunakan bantuan program *web interface*. Model teori respon butir menghasilkan parameter butir yang independen dari peserta tes dan parameter peserta tes yang bersifat independen dari sekumpulan butir yang diujikan (Xitao, 1998). Sebagaimana halnya dalam teori klasik, dalam teori respon butir juga didasari oleh postulat dasar. Ada dua postulat dasar dari teori tes modern (Hambelton, Swaminathan & Rogers, 1991) yaitu: 1) hasil kerja seseorang peserta tes pada suatu butir soal dapat diprediksikan (atau dijelaskan) dari suatu jenis faktor – faktor yang disebut sifat – sifat laten atau kemampuan ;(2) hubungan antara hasil kerja peserta tes pada suatu butir tes dengan sifat – sifat yang mendasarinya dapat didiskripsikan oleh fungsi yang meningkat yang bersifat monotonik yang disebut dengan fungsi karakteristik butir (*Item Characteristic*

Function atau *Item Characteristic Curve – ICC*). Fungsi ini menjelaskan, jika taraf sifat – sifat (kemampuan) meningkat, maka probabilitas menjawab benar pada suatu butir tes juga meningkat.

Dalam *IRT*, kemampuan perorangan tidak dipengaruhi oleh karakteristik item, dan karakteristik item tidak dipengaruhi oleh kemampuan perorangan. Esensi pada *IRT* adalah tingkat kesulitan item dan kemampuan perorangan diukur pada sekla yang sama. Sehingga disini diperlukan kecocokan antara model dengan datanya. *IRT* adalah suatu teori statistik yang membuat model – model matematika yang menyatakan probabilitas respon tertentu terhadap item tertentu sebagai fungsi dari kemampuan seseorang dan karakteristik tertentu dari suatu item (Hambleton, et al, 1991).

Model Logistik yang paling umum adalah dengan tiga (3) Prameter Logistik (PL), yaitu tingkat kesulitan, tebakan, dan diskriminasi. Notasi parameter sedikit berbeda antara buku teks satu dengan yang lain. Model 3 PL ini mengukur tingkat kesulitan (*Difficult*), b_i dan tingkat tebakan (*Gussing*), c_i dan Daya Pembeda (*Discrmn*), a_i peritem. Bentuk umum 3 parameter.

$$P(\theta_i) = c_i + (1 - c_i) \frac{1}{1 + e^{-a_i(\theta - b_i)}} + \epsilon_i \quad (2.3)$$

dengan a_i parameter diskriminasi (*discrimination parameter*), b_i parameter kesulitan (*Difficult*), c_i parameter tebakan (*Gussing*) dan θ tingkat kesulitan. Grafik yang menggambarkan hubungan antara kemampuan dengan peluang menjawab benar disebut grafik ICC (Baker,2001). Untuk $c = 0$ model 3 parameter berubah menjadi 2 parameter, sedangkan jika $c = 0$ dan $a_i = k$ model berubah menjadi 1 parameter (RASCH).

$$P(\theta_i) = c_i + (1 - c_i) \frac{1}{1 + e^{-k(\theta - b_i)}} + \epsilon_i \quad (2.4)$$

Dengan a_i parameter diskriminasi, b_i parameter kesulitan (*Difficult*), c_i parameter tebakan (*Gussing*) dan θ tingkat kesulitan. Grafik yang menggambarkan hubungan antara kemampuan dengan peluang menjawab benar

disebut grafik ICC (*Item Characteristic Curve* Baker,2001). Untuk $a = k$ model 2 parameter berubah menjadi model 1 parameter (RASCH).

Secara kusus Rizopoulos (2006) menggunakan bentuk (notasi disesuaikan)

$$P(x_{im} = 1|z_m) = c_i + (1 - c_i) g (a_i(z_m - b_i)) \quad (2.5)$$

Peluang siswa dengan tingkat kemampuan z_m menjawab benar item ke i yang memiliki parameter kesulitan b_i , parameter diskriminasi a_i , dan parameter tebakan c_i . Fungsi g adalah link untuk model logistik, yaitu logit atau probit. Untuk link logit.

$$g (a_i(z_m - b_i)) = \frac{1}{1+e^{-a_i(z_m-b_i)}} \quad (2.6)$$

Sebagaimana biasanya model (model fitting ada ukuran kecocokan model pemeriksaan model dilakukan dengan menggunakan kriteria informasi Akaike (AIC) yang menghitung perimbangan antara besarnya likelihood dengan banyaknya variabel dengan model. Besarnya AIC dihitung dengan rumus.

$$AIC = -2l(\theta) + 2q, \quad (2.7)$$

Kriteria Item

Kriteria soal (item) yang baik diberikan berikut ini (Anggreyani.2009)

Tabel 2.3 Kriteria Item

No	Aspek	Item Baik
1	Tingkat Kesulitan(dffct b)	$-2 \leq b \leq 2$
2	Daya Pembeda (dscrmn $a = k$)	$a \geq 0.5$
3	Peluang Tebakan (Gussng c)	$c \geq 0.2$

2.4 Item Response Theory (Unidimensi)

Dalam teori respon butir, model matematisnya mempunyai makna bahwa probabilitas subjek untuk menjawab butir dengan benar tergantung pada kemampuan subjek dan karakteristik butir. Ini berarti bahwa peserta tes dengan kemampuan tinggi akan mempunyai probabilitas menjawab benar lebih besar jika dibandingkan dengan peserta yang mempunyai kemampuan rendah. Menurut (Hambleton & Swaminathan 1985) dan (Hambleton, Swaminathan & Rogers 1991) menyatakan bahwa ada tiga asumsi yang mendasari teori respon butir, yaitu unidimensional, independensi lokal dan invariasi parameter. Unidimensi, artinya setiap butir tes hanya mengukur satu kemampuan. Contohnya, pada tes prestasi belajar bidang studi matematika, butir soal yang termuat didalamnya hanya mengukur kemampuan siswa dalam bidang studi matematika saja, bukan bidang yang lain. Dalam pendidikan, siswa menjawab butir soal suatu tes yang berbentuk pilihan ganda dengan benar, biasanya diberi skor 1 dan 0 jika menjawab salah. Pada penyekoran dengan pendekatan tes klasik, kemampuan siswa dinyatakan dengan skor total yang diperolehnya. Prosedur ini kurang memperhatikan interaksi antara setiap orang siswa dengan butir.

Pendekatan teori respon butir merupakan pendekatan alternatif yang dapat digunakan dalam menganalisis suatu tes. Ada dua prinsip yang digunakan pada pendekatan ini, yakni prinsip relativitas dan prinsip probabilitas. Pada prinsip relativitas, unit dasar dari pengukuran bukanlah siswa atau butir, tetapi lebih kepada performance siswa terhadap butir. Jika β_n merupakan indeks dari kemampuan siswa ke n pada trait yang diukur, dan δ_i merupakan indeks dari tingkat kesulitan dari butir ke- i relatif yang terkait dengan kemampuan yang diukur, maka bukan β_n atau δ_i yang merupakan unit pengukuran, tetapi lebih kepada perbedaan antara kemampuan dari siswa relative terhadap tingkat kesulitan butir atau $(\beta_n - \delta_i)$ perlu dipertimbangkan. Sebagai alternatifnya perbandingan antara kemampuan terhadap tingkat kesulitan dapat digunakan. Jika kemampuan dari siswa melampaui tingkat kesulitan butir, maka respons siswa diharapkan benar, dan jika kemampuan siswa kurang dari tingkat kesulitan butir, maka respon siswa diharapkan salah (Keeves dan Alagumalai, 1999).

Pada teori respons butir, prinsip probabilitas menjadi perhatian. Misalkan kemampuan siswa ke n dinyatakan dengan θ_n dan tingkat kesulitan dari butir dinyatakan dengan Δ_i maka sesuai dengan prinsip relativitas, jika $\theta_n > \Delta_i$ siswa diharapkan menjawab dengan benar, dan $\theta_n < \Delta_i$ siswa menjawab salah. Lebih jauh lagi, jika kemungkinan (odds) dari respons siswa terhadap butir diberikan oleh $\frac{\theta_n}{\Delta_i} > 1$ (siswa diharapkan menjawab dengan benar), $\frac{\theta_n}{\Delta_i} < 1$ siswa diharapkan menjawab salah, dan $\frac{\theta_n}{\Delta_i} = 1$, akan terjadi jika kesempatan 50% menjawab benar. Jika peluang menjawab benar, maka $1 - P_{ni}$ merupakan respons tidak benar dan odds untuk respons diberikan oleh :

$$\frac{\theta_n}{\Delta_i} = \frac{P_{ni}}{1-P_{ni}} \text{ dan } \frac{\theta_n}{\Delta_i} = 1 \text{ terjadi jika } P_{ni} = 0,5 \quad (2.8)$$

Probabilitas respons menjawab benar berada pada rentang 0 sampai 1 dan hal ini menghalangi data dinyatakan sebagai skala interval. Sekor mentah yang dihasilkan dari cara ini sulit dinyatakan sebagai skala. Untuk mengatasi permasalahan ini, dapat digunakan transformasi logistik, yang melibatkan logaritma natural dari odds.

$$\ln \frac{\theta_n}{\Delta_i} = \ln \left(\frac{P_{ni}}{1-P_{ni}} \right) \quad (2.9)$$

yang senilai dari

$$\ln \theta_n = \ln \Delta_i = \ln \left(\frac{P_{ni}}{1-P_{ni}} \right) \quad (2.10)$$

Bentuk persamaan yang lebih dikenal dalam pengukuran untuk model ini, yang bisa disebut dengan model Rasch (Hambleton, et al (1991)) :

$$P_i(\theta) = \frac{e^{(\theta-b_i)}}{1+e^{(\theta-b_i)}}, \text{ dengan } i : 1,2,3,\dots,n \quad (2.11)$$

$P_i(\theta)$: probabilitas peserta tes yang memiliki kemampuan θ dipilih secara acak dapat menjawab butir i dengan benar.

θ : tingkat kemampuan subyek (sebagai variabel bebas)

b_i : indeks kesukaran butir ke i

e : bilangan natural yang nilainya mendekati 2,718

n : banyaknya butir dalam tes

2.5 MIRT

Seperti halnya teori respon butir unidimensi, pada model teori respon butir multidimensi data dapat berupa butir skor dikotomi atau politomi. Matriks data disusun sedemikian rupa, dengan x_{ij} menyatakan elemennya pada baris ke i dan kolom ke j . Butir dinyatakan dalam i ($i = 1, \dots, n$) dan peserta dinyatakan dalam j ($j = 1, \dots, n$).

Dalam menyusun matriks data, ada asumsi yang harus diperhatikan (Reckase, 1997), yakni :

- a. Semakin tinggi kemampuan peserta tes, semakin besar probabilitas menjawab benar peserta tes terhadap butir soal (asumsi kemonotonan).
- b. Fungsi probabilitas menjawab benar bersifat smooth (turunan fungsinya terdefiniskan).
- c. Probabilitas kombinasi respons dapat ditentukan dengan hasil probabilitas respons individual ketika probabilitas dihitung kondisional pada titik dalam ruang yang didefinisikan oleh konstruk hipotetik (asumsi independen lokal). Berdasarkan hal ini, asumsi yang digunakan untuk menyusun matriks data yakni asumsi kemonotonan, memiliki turunan fungsi, dan independensi lokal.

Pada teori respons butir multidimensi (*Multidimensional Item Response Theory*) dikenal dua model, yakni *compensatory* dan *noncompensatory*. Menurut Ansley dan Forsyth (Spray, et al., 1990), model *compensatory* membolehkan kemampuan tinggi pada salah satu dimensi memperoleh kompensasi pada kemampuan rendah pada dimensi lain dalam kaitannya dengan probabilitas menjawab benar. Dan berlaku sebaliknya pada model *noncompensatory*. Untuk model *compensatory* pada kasus butir dua dimensi, seorang peserta tes dengan kemampuan sangat rendah pada satu dimensi dan kemampuan sangat tinggi pada dimensi lain dapat menjawab butir tes dengan benar.

Ada dua tipe model *compensatory*, yakni model *MIRT* logistik (Reckase, 1997) dan model ogive normal dari (Samejima) dengan menyatakan kombinasi linier dari kemampuan multidimensi dalam perangkat dalam rumus probabilitas menjawab benar. Dalam model linier ini, rendahnya satu atau lebih kemampuan

dapat dikompensasikan pada dimensi lainnya. Karena kompensasi merupakan karakteristik kombinasi linier, maka model ini diberi nama dengan model *MIRT* linier (Spray, et al.,1990;Bolt & Lall,2003). Model *MIRT* logistik linier dapat ditulis dengan :

$$P_i(\theta_j) = C_i + (1 - C_i) \frac{e^{[\sum_{m=1}^k f_{ijm}] + d_i}}{(1 - e^{[\sum_{m=1}^k f_{ijm}] + d_i})} \quad (2.12)$$

Keterangan :

$P_i(\theta_j)$: Peluang peserta ke- j dengan kemampuan θ_j menjawab benar butir i

θ_j : vektor kemampuan orang ke j

f_{ijm} : nilai yang besarnya sama dengan $a_{im}\theta_{ijm}$

a_{im} : diskriminasi untuk butir ke - i pada dimensi ke - m

θ_{jm} : elemen ke - m dari vektor kemampuan orang ke - $j(\theta_j)$

C_i : parameter pseudo - guessing butir ke- i

d_i : tingkat kesulitan butir ke - i

2.6. Pengertian DIF dan Pendekatan *HGLM*

Misalnya, sebuah item yang dikerjakan oleh kelompok berjenis kelamin laki-laki dan perempuan yang memiliki kemampuan yang setara, memiliki tingkat kesulitan yang lebih tinggi untuk kelompok perempuan dari pada laki-laki. Ini berarti item tersebut memiliki fungsi yang berbeda untuk jenis kelamin yang berbeda, meskipun kedua kelompok itu memiliki kemampuan yang sama. Dalam pengertian ini, sebuah item memungkinkan memiliki (*DIF*) tetapi dapat dianggap tidak bias, jika item tersebut memang mengukur bagian dari kemampuan yang sesuai dengan tujuan pengukuran, hanya saja salah satu kelompok, meskipun memiliki abilitas yang sama, cenderung memiliki kelemahan dalam bagian tersebut. Dampak merupakan pengaruh dari perbedaan kelompok dalam atribut yang diukur pada hasil tes. Ketika perbedaan antar kelompok terjadi mungkin karena adanya perbedaan kuantitas atribut yang diukur dari kedua kelompok, maka tes atau item tersebut dianggap tidak mengalami *DIF*. Perbedaan skor dua subjek yang berasal dari dua kelompok yang berbeda dapat terjadi karena memang

keduanya memiliki kuantitas atribut yang berbeda atau karena item mengalami *DIF*.

Item yang baik seharusnya berlaku adalah bagi setiap peserta tes, tidak boleh menguntungkan suatu kelompok tertentu seperti menguntungkan kelompok perempuan, sehingga merugikan kaum laki – laki. Faktor – faktor yang mempengaruhi *DIF* dibedakan atas faktor internal dan faktor eksternal. Faktor internal adalah faktor – faktor yang berhubungan dengan kondisi responden. Pendeteksian *DIF* dilakukan untuk memeriksa apakah setiap item adil bagi setiap orang tanpa disebabkan karena ada perbedaan bawaan dan responden, seperti perbedaan jenis kelamin, bahasa, suku bangsa, pendidikan orang tua dan lain – lain. Penting untuk mengadakan uji *DIF* pada sebuah alat ukur. Mengingat masyarakat sudah saatnya menjadi lebih kritis dan peduli dengan hasil pengukuran, dan jika pengukuran tersebut dapat dikategorikan gagal atau berhasil dalam sebuah tes. Ada beberapa metode yang dapat digunakan untuk mendeteksi *DIF* Seperti metode menggambarkan ICC, Metode Mantell Haenzel menggunakan 1 parameter, yaitu parameter tingkat kesukaran (Osterlind, 2009).

Model ini akan dijelaskan untuk konteks tanggapan item, dengan siswa sebagai kelompok, item untuk pengamatan berulang, dan respon biner. Data dilambangkan sebagai $Y_{pi} = 0,1$ dengan $p = 1, \dots, P$ sebagai indeks untuk siswa, dan $i = 1, \dots, I$ seperti dalam indeks untuk item. Penggunaan ligan untuk *IRT* terbatas pada data biner dan data kategori pesanan yang dapat didekomposisi menjadi data biner, seperti untuk model rasio kelanjutan (Tutz,1990).

Untuk mengetahui data *GLMM* dapat diperjelas dalam urutan tiga langkah berikut :

- a. Komponen linier: untuk setiap pasangan x dan satu item i (x, i), kombinasi linear dari prediktor menentukan nilai komponen linier. nilai ini dilambangkan disini sebagai (η_{xi}) ,
- b. Menghubungkan komponen: hasil (η_{pi}) , dipetakan ke dalam interval $[0,1]$ berdasarkan fungsi link yang menghasilkan probabilitas (π_{xi}) , nilai yang diharapkan dari (Y_{xi})

- c. Komponen acak: probabilitas (π_{xi}), adalah parameter distribusi Bernoulli / binomial berdasarkan pengamatan biner yang dihasilkan untuk pasangan (x, i) , dilambangkan sebagai $(Y_{xi} \in 0,1)$.

Langkah – langkah GLMM, akan dijelaskan secara rinci dengan urutan terbalik. *Linking component* adalah distribusi Bernoulli. Probabilitas (π_{xi}), adalah parameter distribusi Bernoulli: $Y_{pi} \sim \text{Bernoulli}(\pi_{xi})$. Distribusi Bernoulli adalah distribusi binomial dengan satu pengamatan ($n = 1$). Hal ini khas untuk data *IRT* hanya memiliki satu pengamatan per pasangan dari seseorang dan item sehingga.

$$Y_{xi} \sim \text{Bernoulli}(1, \pi_{xi})$$

$$\beta \binom{n}{1} = p (1 - x)^{1-x} \quad (2.13)$$

Komponen yang menghubungkan memetakan nilai yang diharapkan dari Y_{pi} yaitu (π_{pi}) untuk distribusi Bernoulli pada garis real dari $(-\infty$ sampai $+\infty)$ melalui fungsi link:

$$\eta_{pi} = f_{link}(\pi_{pi}) \quad (2.14)$$

Fungsi link biasanya memilih menjadi fungsi logit, atau $\eta_{pi} = \ln\left(\frac{\pi_{pi}}{1-\pi_{pi}}\right)$. link logit adalah link alami untuk distribusi Bernoulli / binomial, ini mengarah ke model IRT logistik. Alternatif untuk link logit adalah link probit, yang didasarkan pada distribusi normal kumulatif normal, yang juga disebut ogive normal. Kebalikan dari fungsi probabilitas kumulatif menghasilkan nilai fungsi probit link. Kedua link mengarah ke skala yang berbeda. Parameter dalam model probit diskalakan relatif terhadap distribusi normal standar. karena distribusi logistik didekati dengan cukup baik oleh distribusi normal dengan standar deviasi 1,7, skala probit dapat ditransformasikan ke skala logit dengan perkalian dengan $D = 1.7$ yang terkenal. Komponen linear adalah kombinasi linear dari prediktor. Dalam literatur statistik, sering istilah "kovariat" digunakan sebagai pengganti 'prediktor',

dan bobotnya disebut 'efek'. Memperluas dan mengembangkan komponen linier dengan cara yang berbeda merupakan bagian penting dari makalah ini.

Komponen Linier Rasch Model

Komponen linier dari model rasch menggunakan $I + I$ item kovarian. I indikator item ditambah 1 konstan untuk semua item. Sebagai hasil matrik $I + (I + I)$ matrik X dari item kovariat adalah gabungan dari vektor 1 dan matrik identitas $I + I$. Vektor 1 memiliki vektor acak, yang sering disebut kemampuan atau efek laten dalam model IRT. Sedangkan evek dari kovariant dari matrik identitas adalah tetap satu per item, dan pengaruhnya model dapat dirumuskan sebagai berikut:

$$\eta_{pi} = \theta_p X_{i0} + \sum_{k=1}^K \beta_i X_{ik}, \quad (2.15)$$

Dengan $X_{i0} = 1$ untuk semua item; $X_{ik} = 1$ jika $i = k$ ($k = 1, \dots, K$; indeks k memiliki jangkauan yang sama dengan i) dan 0 jika tidak; dan $\theta_p \sim N(0, \sigma_\theta^2)$. Model pada rumus (2.10) juga bisa ditulis lebih sederhana lagu: $\eta_{pi} = \eta_p + \beta_i$. Mengikuti pada 1PL dan equivalen ogive normalnya adalah model penangkap acak. Perhatikan bahwa tanda tambah pada rumus di (2.15) menyiratkan bahwa harus ditafsirkan sebagai kesulitan item. Sebagai prinsip umum, saat melihat model *IRT* dari perspektif *GLMM* semua bentuk pengukuran bergantung pada evek kovariant. Sebagai contoh kebanyakan sifat laten adalah efek acak dan kesulitan item adalah efek tetap.

A. Pendekatan Menggunakan *HGML*

Selain metode *IRT* tradisional untuk mendeteksi *DIF*, beberapa temuan terbaru menggunakan metode alternatif untuk mendeteksi *DIF*. *HGLM* memodelkan data hirarkis. Bila hasilnya adalah data kategoris, seperti data nominal atau ordinal (Raudenbush dan Bryk 2002). Model ini merupakan perluasan General Linier Model (GLM) ke data bertingkat (McCullagh dan Nelder 1989; Kamata 1998). Model level – 1 di *HGML* dua tingkat terdiri dari model sampling, sebuah link fungsi, dan model struktural. Menurut Raudenbush

dan Bryk (2002), Binomial sampling dan logit link digunakan saat hasilnya biner. Berbasis pada distribusi binomial, nilai yang diharapkan dan varians Y_{ij} untuk level 1 Model sampling untuk model 2 dapat ditulis sebagai berikut :

$$E(Y_{ij}|\varphi_{ij}) = \varphi_{ij}, \text{ dan } Var(Y_{ij}|\varphi_{ij}) = \varphi_{ij}(1 - \varphi_{ij})$$

Dimana φ_{ij} adalah probabilitas pemeriksaan yang memberikan respon yang benar terhadap item i , level 1 fungsi logit link di HGML dapat ditulis sebagai berikut :

$$\eta_{ij} = \log\left(\frac{\varphi_{ij}}{1-\varphi_{ij}}\right), \quad \eta_{ij} = \log\left(\frac{\varphi_{ij}}{1-\varphi_{ij}}\right), \quad (2.16)$$

Dimana η_{ij} mewakili log dari peluang pemeriksaan yang memberikan respons yang benar item i , ini bisa mengambil nilai sebenarnya. φ_{ij} dibatasi pada ilai antara 0 dan 1, karena itu adalah probabilitas. Jika φ_{ij} sama dengan 0,5, kemungkinan respon yang benar sama ke 1, yaitu $0,5/0,5 = 1$, dan logit adalah 0, $\log(1) = 0$. Jika φ_{ij} lebih kecil dari 0,5, maka logit itu negatif; jika φ_{ij} lebih besar dari 0,5, maka logit positif .

$$\eta_{ij} = \beta_{0j} + \beta_{1j}X_{1ij} + \beta_{2j}X_{2ij} + \dots + \beta_{kj}X_{kij} = \beta_{0j} + \sum_{h=1}^k \beta_{hj}X_{hij}, \quad (2.17)$$

Dimana X_{hij} adalah variabel dummy indikator – item ke – h untuk teste j , dengan nilai 1 ketika $h = i$ dan 0, ketika $h \neq i$, untuk item i berhubungan dengan koefisien β_{hj} , dimana $h = 1, \dots, k$.

β_{0j} adalah mencegah dari persamaan struktural. Untuk model *HGLM* dua level, yang secara aljabar sama dengan Rasch model, kerangka kerja GLM dan HLM dapat menghasilkan model Rasch (Kamata 2001, 2002). Fungsi link dan model prediktor linier dapat ditentukan sebagai level terendah atau model level item. Sebenarnya, *HGLM* dua tingkat bisa jadi mudah diperluas ke model regresi laten tiga level item, yang diperkirakan akan diestimasi kemampuan kelompok, seperti disekolah dan menunjukkan variasi dari efek interaktif variabel orang dan kelompok (Kamata 2001).

Variabel hasil yang harus dipertimbangkan dikotomis dan diasumsikan mengikuti distribusi binomial (Raudenbush dan Bryk 1986). Link logit fungsi untuk model Rasch tingkat tiga diberikan sebagai (Kamata 2002):

$$\eta_{ijm} = \log \left[\frac{\varphi_{ijm}}{1 - \varphi_{ijm}} \right] \quad (2.18)$$

Dimana φ_{ijm} , adalah probabilitas bahwa respon ke- i benar untuk siswa j disekolah m . η_{ijm} adalah log probabilitas bahwa respon ke- i benar untuk siswa j sekolah m . Bila model Rasch sesuai dengan kerangka *HGLM* dua level, maka persamaan tanpa prediktor level 1 diberikan sebagai

$$\log \left(\frac{\varphi_{ij.}}{1 - \varphi_{ij.}} \right) = \eta_{ij.} = \theta_j - b_i = \theta_j + \beta_i, \quad (2.19)$$

Ketika $\beta_i = -b_i$ dari predictor atau covariant, kondisional level 3 HGML model level 1 dapat ditulis :

$$\log \left(\frac{\varphi_{ij.}}{1 - \varphi_{ij.}} \right) = \eta_{ijm} = \beta_{0jm} + \beta_{1jm}X_{1ijm} + \beta_{2jm}X_{2ijm} + \dots + \beta_{kjm}X_{kijm} = \beta_{0jm} + \sum_{h=1}^k \beta_{hjm}X_{hijm}, \quad (2.20)$$

Dimana variabel hasil dihubungkan ke prediktor dengan fungsi link logistik. X_{hijm} , yang digunakan untuk mengidentifikasi item dalam model prediksi linier, adalah h - th variabel dummy item - indikator ($h = 1, \dots, k$), dengan nilai 1 atau 0, untuk siswa j untuk item i di sekolah m . Koefisien β_{0jm} , adalah istilah yang membatasi dan satu satunya item yang memiliki efek ketika setiap. X_{hijm} , memiliki nilai nol. Koefisien β_{hjm} , adalah terkait dengan X_{hijm} . Model tingkat 1 memprediksi probabilitas siswa j di sekolah m menjawab item soal dengan benar. Model level 2 adalah model tingkat siswa untuk siswa j di sekolah m .

B. Level 2 Model

$$\left\{ \begin{array}{l} \beta_{0jm} = u_{0jm}, \gamma_{00m} + \text{with } u_{0jm} \sim N(0, \tau_\gamma) \\ \beta_{1jm} = u_{10m}, \gamma_{11m}(\text{grup})_{jm} \\ \vdots \\ \beta_{kjm} = u_{k0m}, \gamma_{k1m}(\text{grup})_{jm}, \end{array} \right. \quad (2.21)$$

Dimana (Kelompok) jm adalah variabel indikator untuk keanggotaan kelompok biner, dengan nilai 0 = kelompok focus dan nilai 1 = kelompok referensi. γ_{00m} adalah batas acak mewakili kemampuan rata – rata kelompok tertentu m , menentukan berapa banyak kemampuan siswa berasal, kemampuan siswa di sekolah.

2.7 Struktur Pemrograman Simulasi R-Shiny

Secara umum, komponen program Shiny dibedakan menjadi dua kelompok besar yaitu seperti berikut ini, menurut (Tirta, 2014) :

- a. User Interface. Bagian ini yang bermanfaat untuk

Panel Kontrol adalah panel untuk mengontrol input berupa data, variabel, model, tergantung kompleksitas modul. Tampilan kontrol dapat berupa slider, radio button, check – bok dan lain – lain. Berbagai contoh tampilan kontrol (seperti checkbox, radiobutton, slider dan lain – lain) bisa dilihat pada Rstudio.

Pemasukan permintaan nilai input (data dengan berbagai jenis variabel yang diperlukan, memilih model, jenis dan kriteria uji statistika).

Penyajian Output terkait hasil analisis / uji. Hasil output dapat berupa:

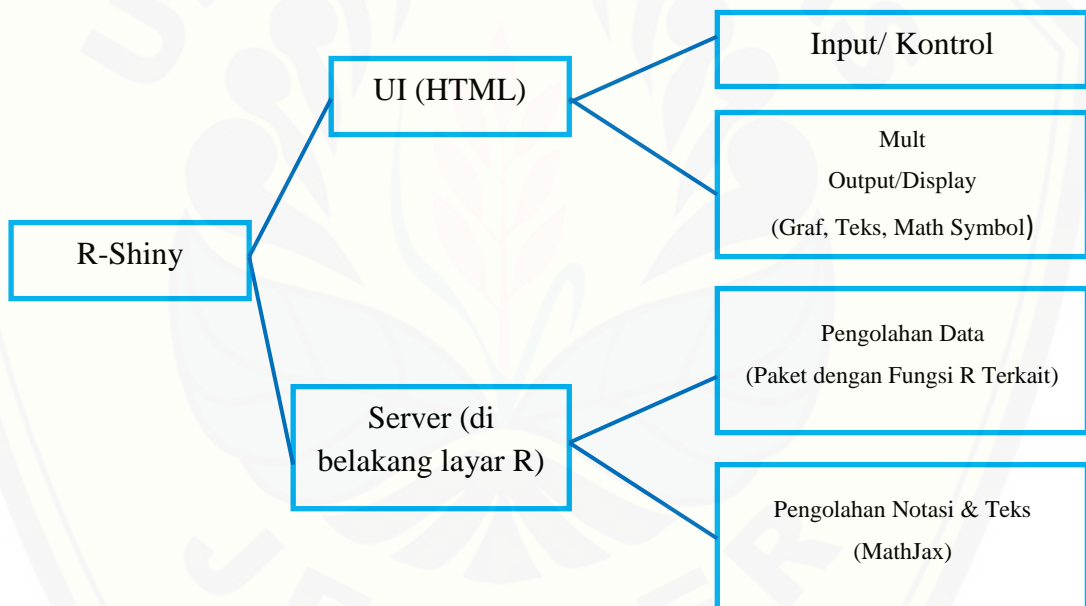
- i. Grafik (histogram, diagram pencar, dan lain – lain)
- ii. Bentuk angka / teks bisa berbentuk asli (verbatim) maupun dalam bentuk tabel.
- iii. Teks khusus dengan notasi matematika dengan format LaTeX

Untuk mengakomodasi berbagai jenis luaran tadi, shiny juga menyediakan berbagai format output seperti plot Output, Verbatim

TexOutput, Tabel Output dan lain – lainnya. Bagian *user interface* ini dapat disajikan pada file khusus ui.r, dapat juga disajikan penuh melalui file HTML, misalnya index.html.

- b. Server. Bagian ini merupakan otak dari program yang bertugas melakukan simulasi , berbagai analisis data sesuai pilihan pengguna dan selanjutnya mengirim hasilnya ke bagian output, bagian ini didukung oleh berbagai prosedur dan analisis data yang pada umumnya telah tersedia pada berbagai paket R. Bagian ini disimpan dalam file yang diberi nama server.r.

Struktur umum pemrograman dengan R- Shiny dapat dilihat pada gambar 2.7



Gambar 2.7. Struktur Komponen R – Shiny

BAB 3. METODE PENELITIAN

3.1 Data Penelitian

Adapun data penelitian ini menggunakan data yang didapat dari 3 lembaga sekolah negeri dan swasta yang menjadi bahan uji tes bank soal. Data yang digunakan berupa nilai akademik dan nonakademik, asal sekolah, jenis kelamin dan juga latar belakang sekolah.

3.2 Tahap Penelitian

Pada tahap ini peneliti menggunakan dua tahap penelitian yaitu menggunakan *web interface* dan penerapannya pada bank soal MGMP.

3.2.1 Penyusunan *Web Interface*

Menyesuaikan (*GLMM*). Kedua efek tetap dan efek acak ditentukan melalui formula model.

Usage

```
glmer(formula, data = NULL, family = gaussian, control =  
glmerControl(),  
      start = NULL, verbose = 0L, nAGQ = 1L, subset, weights,  
na.action,  
      offset, contrasts = NULL, mustart, etastart,  
devFunOnly = FALSE, ...)
```

Formula diatas berfungsi khusus untuk mixed model, perintah untuk sintaks seperti variabel, item random efek, dan dif, yang selanjutnya diubah atau disajikan dalam tampilan bentuk menu yang ramah bagi pengguna web.

Secara konstruktif, *Plmix*, dalam banyak hal, ciri laten atau model *IRT*, dikembangkan dalam pengujian pendidikan dan literatur psikometrik, mewakili beberapa *GLMM* terdahulu. *MIRT* memperkirakan data respon *dikotomis* dan *polytomous* menggunakan model laten unidimensional dan multidimensi berdasarkan paradigma *IRT* ini diperkirakan bebas. Untuk melakukannya, kita harus menentukan matriks *lambda*, yang akan berisi struktur faktor. Untuk contoh ini, kita memerlukan pemuatan terpisah untuk setiap faktor pada setiap titik waktu. kita dapat mengumpulkan poin waktu. Sesuaikan model efek campuran linier (*GLMM*) linier dengan struktur faktor. Gunakan kedua paket *lme4* dan fungsi optimis untuk estimasi menggunakan pendekatan berbasis profil.

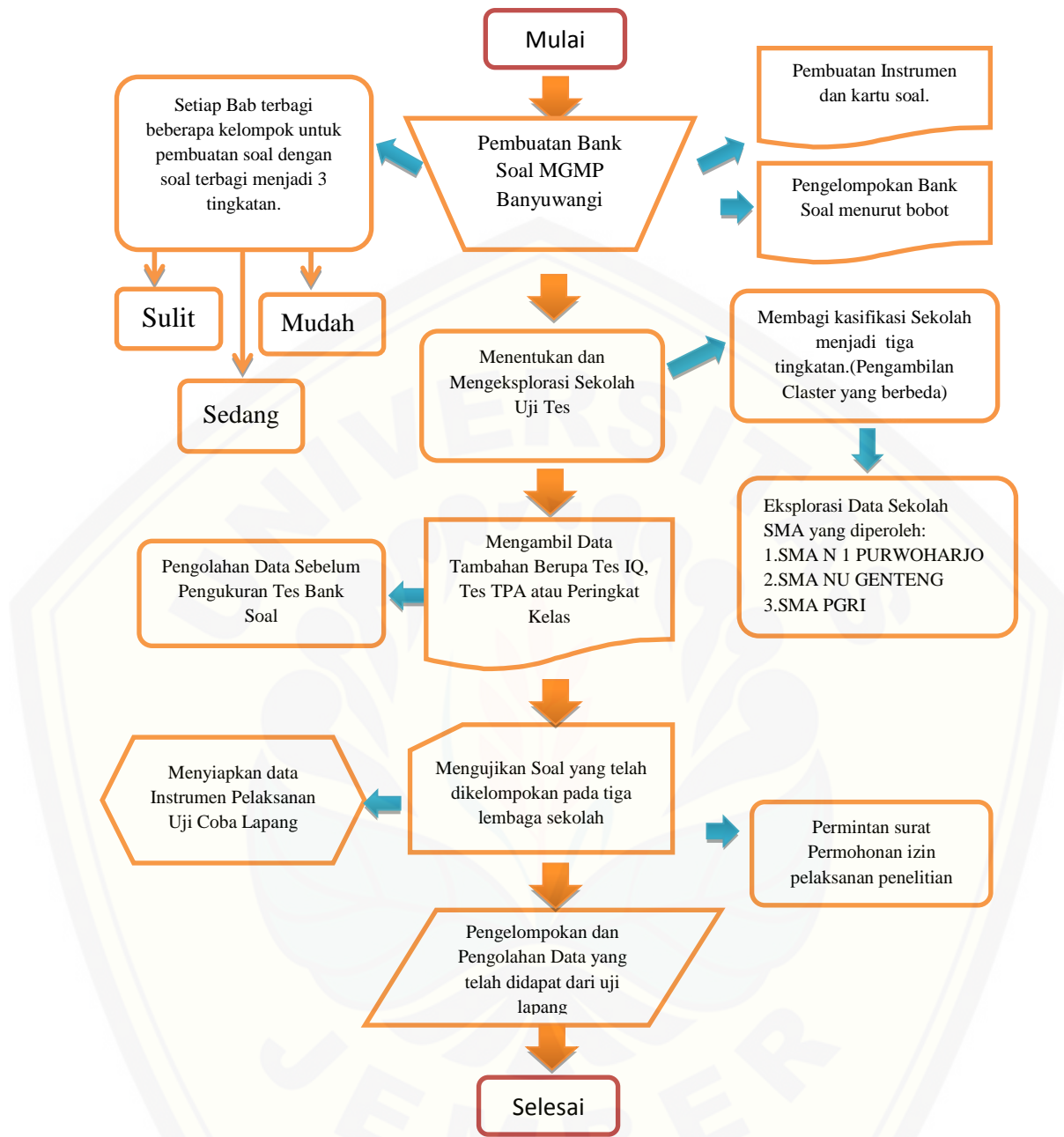
Usage

```
PLmixed(formula, data, family = gaussian, load.var = NULL, lambda = NULL, factor = NULL, init = 1, nlp = NULL, init.nlp = 1, nAGQ = 1, method = "L-BFGS-B", lower = -Inf, upper = Inf, lme4.optimizer = "bobyqa", lme4.start = NULL, lme4.optCtrl = list(), opt.control = NULL, REML = FALSE, SE = 1, ND.method = "simple", est = TRUE)
```

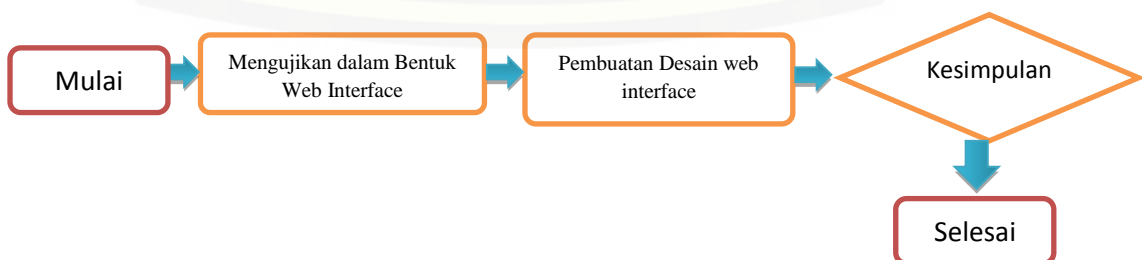
3.2.2. Penerapan Data Bank Soal

Penelitian dilaksanakan di universitas Jember fakultas FMIPA dengan mengikuti proses tahap penelitian pada Gambar 3.1 secara terperinci. Uji analisis validasi bank soal MGMP di LabMatematika FMIPA Universitas Jember. Tahap proses penelitian terbagi menjadi lima tahapan pokok inti yaitu :

- a. Kegiatan yang diawali dengan kegiatan observasi, persiapan meliputi pembuatan bank soal dengan mengikuti MGMP Banyuwangi. Penyiapan instrumen bank soal dilakukan dengan cara membuat kelompok dalam setiap sub bab topik yang didalamnya terdiri dari 3 jenis soal dengan tingkat kesukaran yang berbeda dimulai dari mudah, sedang dan sulit.
- b. Pengumpulan claster Sekolah uji coba dan pemilahan soal sesuai kelompok topik validasi soal. Selain itu pada tahap ini peneliti membutuhkan data tambahan berupa tes IQ, TPA atau rengking siswa pada terakhir semester. Digunakan sebagai bahan tambahan data sebelum uji validasi soal pada siswa. Khususnya pada siswa kelas XII IPA.
- c. Eksplorasi dengan mengujikan validasi soal pada tiga sekolah yang memiliki claster yang berbeda. Dengan menguji bank soaldilapangan pada siswa kelas XII.
- d. Analisis Data, yang meliputi analisis data dari hasil pengukuran pada siswa yang memiliki claster yang berbeda. Dengan cara pengumpulan dan melihat hasil yang dicapai siswa dengan membandingkan dari data sebelum uji coba tes.
- e. Menampilkan data pada web interaktif dengan menyajikan hasil dari data yang diperoleh pada saat eksplorasi dilakukan pada 3 lembaga sekolah.



Gambar 3.1 Diagram alir tahap penelitian pelaksanaan uji bank soal MGMP Banyuwang.



Gambar 3.2 Diagram alir tahap pengujian dalam bentuk *Web Interface*

3.3. Observasi Awal

Pada penelitian ini menggunakan data yang didapat dari observasi awal pada tiga lembaga sekolah yang memiliki klaster yang berbeda untuk membandingkannya namun dengan batasan kelas XII IPA saja yang menjadi objek penelitian dalam tesis ini sebagai bahan data awal, yaitu dari SMA PGRI Tegaldlimo, SMA NU Genteng dan SMA Negeri 1 Purwoharjo. Peneliti memilih sekolah yang memiliki lingkungan yang berbeda dalam setiap sekolah. Masing – masing dari sekolah ini memiliki tingkat yang berbeda dari tingkat kemampuan akademik dan juga fasilitas sekolah yang menunjang adanya kegiatan pembelajaran yang maksimal.



Gambar 3.3 Tahapan Proses dan Progress Kegiatan Penelitian

Pada gambar 3.3 Tahapan Progres dan Progress Kegiatan Penelitian dimulai dengan observasi dan persiapan instrumen yang akan digunakan sebagai informasi. Pada observasi awal juga untuk mendapatkan gambaran data sebelum pelaksanaan tes uji bank soal.

BAB 5. PENUTUP

5.1 Kesimpulan

Peneliti dalam tesis ini telah mengembangkan *Web Interface* untuk Analisis *Item Response Theory* dengan Pendekatan Model Campuran dan Penerapannya pada Bank Soal MGMP. Berdasarkan hasil penelitian dengan menggunakan *Web Interface* yang dikembangkan, menghasilkan beberapa menu diantaranya :

1. *Web Interface* memudahkan dengan tampilan fitur - fitur yang mamprediksi tiap item soal yang memiliki kekurangan atau perlunya evaluasi pada tiap soal yang dinyatakan kurang baik. Serta cara kerja yang mempermudah guru memasukkan banyak data ini menunjukkan web interface lebih efisien dibandingkan dengan yang berbasis Microsoft Excel.
2. Dari hasil analisis uji coba bank soal MGMP di Banyuwangi, terdapat 19 soal yang dinyatakan kurang baik atau adanya revisi pada soal nomor 2, 5, 8, 9, 10, 13, 17, 19, 21, 23, 24, 24, 29, 34, 37, 38, 40, 42 dan 48.
3. Dengan efek nilai bahasa Indonesia dan efek gender, sekolah (sekolah dianggap acak mewakili sekolah) Hasil kontribusi Gender dan Bhs Indonesia tidak sepenting kemampuan untuk menjawab pertanyaan dengan benar. Ini juga berarti tidak bias (DIF) Jenis kelamin meskipun tidak signifikan, jika semuanya dipertahankan dalam model maka diperoleh hasil yang baik.

5.2 Saran

Berdasarkan kesimpulan maka perlu adanya pengembangan untuk model HGLM (*Hierarchical Generalized Linear Mixed Model*) banyak komponen variabel yang mempengaruhi siswa menjawab benar atau salah, karena belum tentu jenis soalnya saja yang baik atau tidak ada faktor lain sehingga siswa menjawab soal benar dan salah pada setiap item butir soal untuk itu diperlukan faktor pendukung yang perlu diperhatikan ketika melakukan uji item butir soal.

DAFTAR PUSTAKA

- Ackerman, T.A., Gierl, M.J., & Walker, C.M. (2003). *Using Multidimensional item response theory to evaluate education and psychological tests*. Education Measurement, Vol.22, pp.37-53.
- Arikunto, S. 1999. *Dasar –dasar Evaluasi Pendidikan*. Jakarta : Bumi Aksara.
- Allen, M.J & W.M Yen. 1979. *Introduction to measurement theory*. Montere: Books/Cole Publising Company.
- Anggreyani, A. 2009. *Penerapan Teori Uji Klasik dan Teori Respon Butir dalam Mengevaluasi Butir Soal* (Skripsi Jurnal Statistika IPB)
- Blejec, A. 2003. *Teaching Statistics by using simulations on the internet. IASE/ISI Satellite*.
- Bolt, D.M & Lall, V.M (2003). *Estimation of compensatory and noncompensatory multidimensional item response models using Marcov chain Monte Carlo. Applied Psychological Measurement*, No. 27, pp. 395 – 414.
- Boroel B. & Gonzalez M. (2017). *Development of a Scale to Measure Attitudes toward Professional Values: An analysis of dimensionality using Rasch measurement*: ScienceDirect; www.elsevier.com/locate/jval.
- De Boeck P et al (2011). *The Estimation of Item Response Models with the Lmer Function from the lme4 Package in R. Journal of Statistical Software*; <http://www.jstatsoft.org/>.
- Djemari, M. (2008). *Teknik penyusunan instrumen tes dan nontes*. Yogyakarta: Mitra Cendekia Press.
- Embretson, E. & Reise, S.P. (2000). *Item response theory for psychologists*. Mahwah: NJ Publications, Lawrence Erlbaum Associates.
- Fahrmeir L, Tutz G (2001). *Multivariate Statistical Modeling Based on Generalized Linear Mixed Models*. 2nd edition. Springer – Verlag, New York.
- Hambleton, R.K & Swaminathan, H. (1985). *Item response theory , principles, and applications*. Boston: Kluwer. Nijhoff Publising.
- Hambleton, R.K & Swaminathan, H. & Rogers, H.J. (1991). *Fundamentals of item response Theory*. California: Sage Publications, The Internasional Professional Publishers.
- Hayat, B. 1997. *Pedoman Penggunaan Iteman*. Jakarta : PUSISJIAN Balitbang Depdiknes.

- Kilmen S & Demirtasli N.(2012). *Coparison of test equating methods based on item response theory according to the sample size and ability distribution*.ScienceDirect; www.elsevier.com/locate/jval.
- Knol, D.L & Berger, MP.F.(1991). *Empirical comparison between factor analysis and multidimensional item response models*. *Multivariate Behavioral Research*, No. 26, pp. 457-477.
- Kose Ibrahim A(2012). *Comparison of unidimensional and multidimensional models based on item response theory in terms of both variables of test legth and sample size*.SciVerse ScienceDirect; www.elsevier.com/locate/jval.
- Lord,M.L. (1980). *Applications of item response theory to practical testing problems*. New Jersey: Lawrence Erlbaum Associates, Publisher.
- Measures:A Comparison of Worked Examples.ScienceDirect; journal homepage: www.elsevier.com/locate/jval.
- Naga, D. S. 1992. Pengantar Teori Sekor pada Pengukuran Pendidikan Jakarta: Besbats.
- Osterlind, S.J., & Everson, H.T. (2009). *Differential Item Functioning*. Second Edition. Thousand Oaks (CA): Sage Publications
- Petrillo Jennifer,(2015).Using Classical Test Theory, Item Response Theory, and Rasch Measurement Theory tonEvaluate Patient – Reported Outcom
- Retnawati H,(2008). Estimasi relative tes berdasarkan teori tes klasik dan teori respons butir. Disertasi. Universitas Negeri Yogyakarta,tidak dipublikasikan.
- Retnawati H (2014). Teori Respons Butir dan Penerapannya,Yogyakarta. Email: nuhamedika@gmail.com – nuhamedika@yahoo.com.
- Reckase, M.D. (1997). *A linier logistic multidimensional model for dicotomous item response data*. In W.J Linden & R.K Hambleton (Eds), *Handbook of modern item response theory* (pp. 271 – 286). New York : Springer.
- Rijmen F, Tuerlinckx F, De Boeck P, Kuppens P (2003). “ A Nonlinear Mixed Model Framework for Item Response Theory.” *Psychological Methods*,8.
- Rizopoulos D.2006. *ltm An R Package for Laten Variabel Modeling and Item Response Theory Analysis*. *Journal of matematical Software* vol.17(5):1-25.
- Rstudio and Inc. Shiny Widget Gallery. <http://shiny.rstudio.com/gallery/widget-gallery.html> [September 2014]
- Rabe- Hesketh dan Jeon Minjeong, 2012 “ *Profile – Likelihood Approach for Estimating Generalized Linear Mixed Models with Factor Strutures*. “ *Jurnal of Education and Behavioral Statistic* 37 (4). SAGE Publications Sage CA: Los Angeles, CA : 518 – 42.

- Sanaky,H.(1998). Teknik menyusun alat evaluasi belajar mata pelajaran Al- Islam dan bahasa Arab. Makalah disajikan dalam acara Pembinaan Guru Madrasah Mu'alimat Muhammadiyah, tanggal 26 September 1998. Diambil 20 November 2017<http://www.docstoc.com/docs/18529273>.
- Sanchez,G.(2013) PLS Path Modeling with R. Trowchez Editions.Berkeley, 2013.
<http://www.gastonsanchez.com/PLS Path Modeling with R.pdf>.
- Suwarto,(2012) Teori Tes Klasik dan Teorites Modern. Jl.Sujono Humardani No.1 Jombang Sukoharjo. E-mail: suwartowarto@yahoo.com .
- Spray,J.A,Davey,T.C.,Rechase,M>D.,et al.(1990). *Comparison of two logistic multidimensional item response theory models. ACT Research Report Series. United States Government.*
- Tirta, IM. (2015). Pengembangan Analisis Respon Item Interaktif Online Menggunakan R Untuk Respon Dikotomus Dengan Model Logistik (1-PI, 2-PI-3PI) *Prosiding Seminar Nasional Pendidikan dan Produk Akademik Universitas Jember* 30 Mei 2015 hal : 420 – 427.
- Tutz G (1990). “*Sequential Item Response Model with an Ordered Response.*” *British Journal of Mathematical and Statistical Psycology*,43,39-55.
- Virtual Laboratorium Statistika Universitas Jember (Pusat Belajar dan Berlatih statistika online dengan R-Shiny (<http://103.241.207.58>))
- Xitao Fan. (1998). Item response theory and classical test theory : an empirical comparison of their item/person statistic. *Educational and Psychological Measurement*, 58,357 – 381.
- Youngchim P(2014). Development of a Mathematical Problem Solving Diagnostic Method: an Application of Bayesian Networks and Multidimensional Item Respond Theory. *ScientDirect*:
www.sciencedirect.com.

Lampiran A: Data Siswa

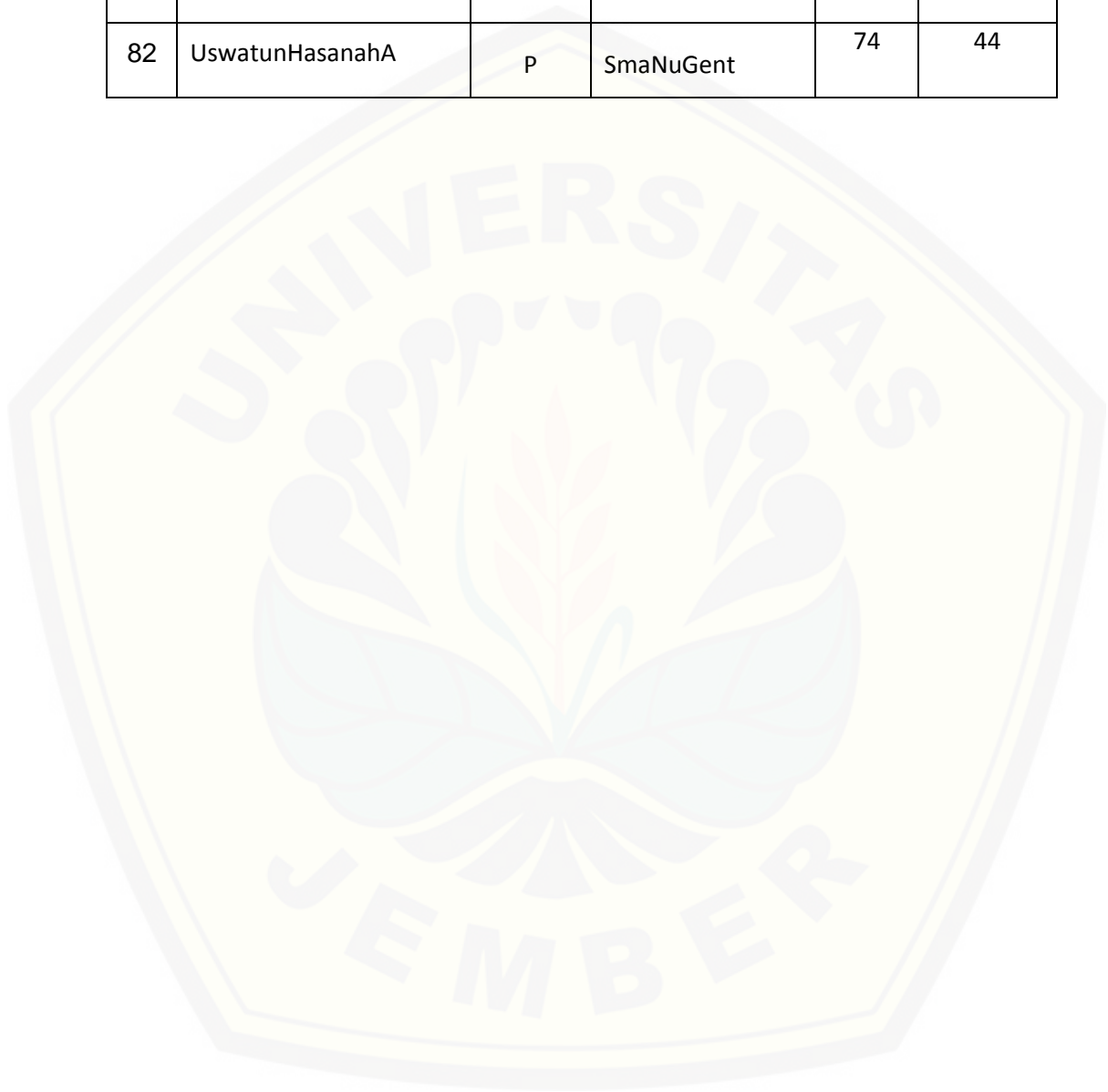
IdS	Nama	Gender	AsalSekolah	NilaiBhs sIndo	Nilai Hasil Tes
1	AdamAnugrahS	L	SmaPgriTeg	67	56
2	AminSetyoadi	L	SmaPgriTeg	68	58
3	AndiSetiawan	L	SmaPgriTeg	62	50
4	ArdianSaputra	L	SmaPgriTeg	64	44
5	ArsyLuxyWahyuni	L	SmaPgriTeg	66	48
6	AsatMiftahAnugrah	L	SmaPgriTeg	67	54
7	DandiRahmaFajar	L	SmaPgriTeg	68	46
8	DityaNikenAriPuspita	P	SmaPgriTeg	78	52
9	EkaPrastyowati	P	SmaPgriTeg	67	34
10	Fitriani	P	SmaPgriTeg	62	42
11	HepiNofitaSari	P	SmaPgriTeg	64	44
12	LeoFannySuwondo	L	SmaPgriTeg	64	48
13	MalaysiaIntanHidayah	P	SmaPgriTeg	72	40
14	MohDianPurnama	L	SmaPgriTeg	78	48
15	MuhVikaSofani	L	SmaPgriTeg	78	40
16	PingkyAdiNugroho	L	SmaPgriTeg	76	50
17	PriyoAgung	L	SmaPgriTeg	62	28
18	RuliPrasetyo	L	SmaPgriTeg	72	38

19	SoniatulLutfia	P	SmaPgriTeg	78	50
20	UswatunKhoiriyah	P	SmaPgriTeg	67	40
21	VinkaIndriyani	P	SmaPgriTeg	68	44
22	ViviSuryaNingsih	P	SmaPgriTeg	76	46
23	WayanSlamet	L	SmaPgriTeg	78	44
24	WuriWrasaptuningtyas	P	SmaPgriTeg	68	46
25	YosiTriBuana	P	SmaPgriTeg	76	52
26	AbeliaFaridhotul	P	SmaNegriPurwo	78	58
27	AgrendiRezaSaputra	L	SmaNegriPurwo	76	58
28	AlberthaAyuningTyas	P	SmaNegriPurwo	78	50
29	Aldy Rahman	L	SmaNegriPurwo	78	44
30	ArsintaQiftiani	P	SmaNegriPurwo	76	48
31	BachtiarFermansyah	L	SmaNegriPurwo	80	54
32	CandraDimasFantauza	L	SmaNegriPurwo	74	46
33	CiciAkmalia	P	SmaNegriPurwo	82	52
34	DelaNastasyaYunita	P	SmaNegriPurwo	78	34
35	DenalhzaMatalata	L	SmaNegriPurwo	80	42
36	DianAnnisa	P	SmaNegriPurwo	76	44
37	DiopientaDhofatulRid	L	SmaNegriPurwo	68	48
38	DwiBagusAgung	L	SmaNegriPurwo	80	40

39	Dwiyankusuma	L	SmaNegriPurwo	80	48
40	EgaYuliaKartikaSari	P	SmaNegriPurwo	76	40
41	FrizkyEkaRamadani	P	SmaNegriPurwo	80	50
42	GusmatWaisAl-Qurni	L	SmaNegriPurwo	80	28
43	HanindyaBerlianaAS	P	SmaNegriPurwo	80	38
44	HilmaUlyaturRosida	L	SmaNegriPurwo	78	50
45	IlhamVerdiyansah	L	SmaNegriPurwo	80	40
46	MartaDiasFelina	P	SmaNegriPurwo	76	44
47	MerryNovitasari	P	SmaNegriPurwo	80	58
48	MohNizar	L	SmaNegriPurwo	80	42
49	NadineAlfiaM	P	SmaNegriPurwo	78	40
50	NovitaNurHamidah	P	SmaNegriPurwo	62	40
51	PuriWulandari	P	SmaNegriPurwo	76	48
52	SindiDwiSaputri	P	SmaNegriPurwo	78	54
53	VelaMartaTrisanti	P	SmaNegriPurwo	80	46
54	VintriDwiRahayu	P	SmaNegriPurwo	76	52
55	ViraAlisaMardilita	P	SmaNegriPurwo	80	34
56	WahyuFalsafa	P	SmaNegriPurwo	80	42
57	WaysAlqorni	L	SmaNegriPurwo	78	44
58	WidyaSuryaning	P	SmaNegriPurwo	80	48

59	YogiBalinda	L	SmaNegriPurwo	76	40
60	ZogiQurotaAyuni	L	SmaNegriPurwo	78	48
61	ZianiRefisa	P	SmaNegriPurwo	72	40
62	AhmadKoimam	L	SmaNuGent	76	58
63	ChidziamChudzaqi	L	SmaNuGent	78	50
64	DwiNuryana	P	SmaNuGent	76	44
65	EndangSriLestari	P	SmaNuGent	74	48
66	KhoirulAnwarAziz	L	SmaNuGent	76	54
67	LinggarMukti	L	SmaNuGent	78	46
68	LusiAristiani	P	SmaNuGent	76	52
69	MiftakulFirdausiA	L	SmaNuGent	76	34
70	MiftakhurRoifah	L	SmaNuGent	62	42
71	MohMahrusBaihaqi	L	SmaNuGent	64	44
72	MohSahrulMaqom	L	SmaNuGent	64	48
73	MohammadYusriLP	L	SmaNuGent	78	40
74	MohammadRamad	L	SmaNuGent	78	48
75	NikmatusSholihah	P	SmaNuGent	74	40
76	Nurlmamah	P	SmaNuGent	74	48
77	NurulKhomariyah	P	SmaNuGent	76	40
78	PutriAyuLestari	P	SmaNuGent	62	28

79	RudyHartono	L	SmaNuGent	74	38
80	UlulMeiNovita	P	SmaNuGent	78	50
81	UmiNadiratulLaili	P	SmaNuGent	76	40
82	UswatunHasanahA	P	SmaNuGent	74	44



Lampiran B : Ringkasan Data pada Interface

I1			
Min. :0.0000			
1st Qu.:0.0000			
Median :1.0000			
Mean :0.6098			
3rd Qu.:1.0000			
Max. :1.0000			
I2	I3	I4	I5
Min. :0.0000	Min. :0.0000	Min. :0.0000	Min. :0.0000
1st Qu.:1.0000	1st Qu.:0.0000	1st Qu.:0.0000	1st Qu.:0.0000
Median :1.0000	Median :0.0000	Median :0.0000	Median :0.0000
Mean :0.7683	Mean :0.4756	Mean :0.2927	Mean :0.2073
3rd Qu.:1.0000	3rd Qu.:1.0000	3rd Qu.:1.0000	3rd Qu.:0.0000
Max. :1.0000	Max. :1.0000	Max. :1.0000	Max. :1.0000
I6	I7	I8	I9
Min. :0.0000	Min. :0.0000	Min. :0.0000	Min. :0.0000
1st Qu.:0.0000	1st Qu.:0.0000	1st Qu.:0.0000	1st Qu.:0.0000
Median :0.0000	Median :0.0000	Median :1.0000	Median :0.0000
Mean :0.4878	Mean :0.3659	Mean :0.5122	Mean :0.4024
3rd Qu.:1.0000	3rd Qu.:1.0000	3rd Qu.:1.0000	3rd Qu.:1.0000
Max. :1.0000	Max. :1.0000	Max. :1.0000	Max. :1.0000
I10	I11	I12	I13
Min. :0.0000	Min. :0.0000	Min. :0.0000	Min. :0.0000
1st Qu.:1.0000	1st Qu.:0.0000	1st Qu.:0.0000	1st Qu.:0.0000
Median :1.0000	Median :0.0000	Median :0.0000	Median :1.0000
Mean :0.7561	Mean :0.4878	Mean :0.2073	Mean :0.7195
3rd Qu.:1.0000	3rd Qu.:1.0000	3rd Qu.:0.0000	3rd Qu.:1.0000
Max. :1.0000	Max. :1.0000	Max. :1.0000	Max. :1.0000
I14	I15	I16	I17
Min. :0.0000	Min. :0.0000	Min. :0.0000	Min. :0.0000
1st Qu.:0.0000	1st Qu.:0.0000	1st Qu.:0.0000	1st Qu.:0.0000
Median :1.0000	Median :0.0000	Median :1.0000	Median :1.0000
Mean :0.6098	Mean :0.4634	Mean :0.5244	Mean :0.6098
3rd Qu.:1.0000	3rd Qu.:1.0000	3rd Qu.:1.0000	3rd Qu.:1.0000
Max. :1.0000	Max. :1.0000	Max. :1.0000	Max. :1.0000
I18	I19	I20	I21

Min. :0.0000	Min. :0.0000	Min. :0.0000	Min. :0.0000
1st Qu.:0.0000	1st Qu.:0.0000	1st Qu.:0.0000	1st Qu.:0.0000
Median :0.0000	Median :0.0000	Median :1.0000	Median :1.0000
Mean :0.4878	Mean :0.1707	Mean :0.5976	Mean :0.6585
3rd Qu.:1.0000	3rd Qu.:0.0000	3rd Qu.:1.0000	3rd Qu.:1.0000
Max. :1.0000	Max. :1.0000	Max. :1.0000	Max. :1.0000
I22	I23	I24	I25
Min. :0.0000	Min. :0.0000	Min. :0.0000	Min. :0.0000
1st Qu.:1.0000	1st Qu.:0.0000	1st Qu.:0.0000	1st Qu.:0.0000
Median :1.0000	Median :0.0000	Median :0.0000	Median :1.0000
Mean :0.8049	Mean :0.1829	Mean :0.2561	Mean :0.5366
3rd Qu.:1.0000	3rd Qu.:0.0000	3rd Qu.:0.7500	3rd Qu.:1.0000
Max. :1.0000	Max. :1.0000	Max. :1.0000	Max. :1.0000
I26	I27	I28	I29
Min. :0.00000	Min. :0.0000	Min. :0.0000	Min. :0.0000
1st Qu.:0.00000	1st Qu.:1.0000	1st Qu.:0.0000	1st Qu.:0.0000
Median :0.00000	Median :1.0000	Median :1.0000	Median :1.0000
Mean :0.04878	Mean :0.7561	Mean :0.7317	Mean :0.7073
3rd Qu.:0.00000	3rd Qu.:1.0000	3rd Qu.:1.0000	3rd Qu.:1.0000
Max. :1.00000	Max. :1.0000	Max. :1.0000	Max. :1.0000
I30	I31	I32	I33
Min. :0.0000	Min. :0.0000	Min. :0.0000	Min. :0.0000
1st Qu.:0.0000	1st Qu.:0.0000	1st Qu.:0.0000	1st Qu.:0.0000
Median :1.0000	Median :1.0000	Median :0.0000	Median :1.0000
Mean :0.5244	Mean :0.6585	Mean :0.4756	Mean :0.6341
3rd Qu.:1.0000	3rd Qu.:1.0000	3rd Qu.:1.0000	3rd Qu.:1.0000
Max. :1.0000	Max. :1.0000	Max. :1.0000	Max. :1.0000
I34	I35	I36	I37
Min. :0.00000	Min. :0.0000	Min. :0.0000	Min. :0.0000
1st Qu.:0.00000	1st Qu.:0.0000	1st Qu.:0.0000	1st Qu.:0.0000
Median :0.00000	Median :0.0000	Median :0.0000	Median :1.0000
Mean :0.03659	Mean :0.2195	Mean :0.439	Mean :0.6341
3rd Qu.:0.00000	3rd Qu.:0.0000	3rd Qu.:1.0000	3rd Qu.:1.0000
Max. :1.00000	Max. :1.0000	Max. :1.0000	Max. :1.0000
I38	I39	I40	I41
Min. :0.0000	Min. :0.0000	Min. :0.0000	Min. :0.0000
1st Qu.:0.0000	1st Qu.:0.0000	1st Qu.:0.0000	1st Qu.:0.0000
Median :1.0000	Median :1.0000	Median :0.0000	Median :0.0000

Mean :0.7317	Mean :0.5488	Mean :0.1585	Mean :0.4756
3rd Qu.:1.0000	3rd Qu.:1.0000	3rd Qu.:0.0000	3rd Qu.:1.0000
Max. :1.0000	Max. :1.0000	Max. :1.0000	Max. :1.0000
I42	I43	I44	I45
Min. :0.0000	Min. :0.0000	Min. :0.0000	Min. :0.0000
1st Qu.:0.0000	1st Qu.:0.0000	1st Qu.:0.0000	1st Qu.:0.0000
Median :0.0000	Median :0.0000	Median :1.0000	Median :0.0000
Mean :0.1951	Mean :0.4756	Mean :0.5366	Mean :0.3659
3rd Qu.:0.0000	3rd Qu.:1.0000	3rd Qu.:1.0000	3rd Qu.:1.0000
Max. :1.0000	Max. :1.0000	Max. :1.0000	Max. :1.0000
I46	I47	I48	I49
Min. :0.0000	Min. :0.0000	Min. :0.0000	Min. :0.0000
1st Qu.:0.0000	1st Qu.:0.0000	1st Qu.:0.0000	1st Qu.:0.0000
Median :1.0000	Median :1.0000	Median :0.0000	Median :0.0000
Mean :0.5366	Mean :0.5732	Mean :0.439	Mean :0.2439
3rd Qu.:1.0000	3rd Qu.:1.0000	3rd Qu.:1.0000	3rd Qu.:0.0000
Max. :1.0000	Max. :1.0000	Max. :1.0000	Max. :1.0000
I50			
Min. :0.0000			
1st Qu.:0.0000			
Median :1.0000			
Mean :0.6707			
3rd Qu.:1.0000			
Max. :1.0000			

Lampiran C: Data Hasil Jawaban Siswa

IdS I12	Gender	AsalSekolah	NilaiBhsIndo	I1	I2	I3	I4	I5	I6	I7	I8	I9	I10	I11
1 0	1 1	L	SmaPgriTeg	67	1	1	0	0	1	1	1	0	1	0
2 0	2 0	L	SmaPgriTeg	68	0	1	0	0	1	0	0	1	0	1
3 1	3 0	L	SmaPgriTeg	62	1	1	0	0	0	1	0	0	0	1
4 0	4 0	L	SmaPgriTeg	64	0	0	0	0	0	0	1	1	1	1
5 1	5 0	P	SmaPgriTeg	66	1	1	1	0	0	1	0	0	0	1
6 1	6 1	L	SmaPgriTeg	67	1	1	1	0	0	1	1	1	1	1
7 1	7 0	L	SmaPgriTeg	68	1	1	1	0	0	0	0	0	0	0
8 0	8 0	P	SmaPgriTeg	78	1	1	0	1	0	1	0	1	1	1
9 0	9 0	P	SmaPgriTeg	67	0	0	0	1	0	0	0	1	1	0
10 1	10 1	P	SmaPgriTeg	62	1	1	1	0	0	0	0	1	0	1
11 0	11 0	P	SmaPgriTeg	64	1	1	1	0	1	0	0	0	1	1
12 1	12 0	L	SmaPgriTeg	64	1	0	1	0	0	1	0	1	0	1
13 1	13 0	P	SmaPgriTeg	72	1	1	0	0	0	1	1	1	0	1
14 0	14 0	L	SmaPgriTeg	78	0	1	1	0	0	1	0	0	1	1
15 0	15 0	L	SmaPgriTeg	78	0	1	1	0	1	0	0	0	0	1
16 1	16 0	L	SmaPgriTeg	76	0	1	1	1	0	1	1	1	1	1
17 1	17 0	L	SmaPgriTeg	74	1	1	0	0	0	0	0	0	0	1
18 0	18 0	L	SmaPgriTeg	72	0	1	0	1	0	0	1	0	0	0
19 1	19 0	P	SmaPgriTeg	78	1	0	1	0	0	1	0	1	0	1

60 1	60 1	L	SmaNegriPurwo	78	1	1	1	0	0	0	0	1	0	1
61 0	61 0	P	SmaNegriPurwo	80	1	1	1	0	1	0	0	0	1	1
62 1	62 0	L	SmaNuGent	76	1	0	1	0	0	1	0	1	0	1
63 1	63 0	L	SmaNuGent	78	1	1	0	0	0	1	1	1	0	1
64 0	64 0	P	SmaNuGent	76	0	1	1	0	0	1	0	0	1	1
65 0	65 0	P	SmaNuGent	74	0	1	1	0	1	0	0	0	0	1
66 1	66 0	L	SmaNuGent	76	0	1	1	1	0	1	1	1	1	1
67 1	67 0	L	SmaNuGent	78	1	1	0	0	0	0	0	0	0	1
68 0	68 0	P	SmaNuGent	76	0	1	0	1	0	0	1	0	0	0
69 1	69 0	L	SmaNuGent	76	1	0	1	0	0	1	0	1	0	1
70 1	70 1	L	SmaNuGent	62	1	0	1	1	0	0	0	1	0	1
71 1	71 1	L	SmaNuGent	64	1	1	1	1	0	1	1	0	0	0
72 0	72 0	L	SmaNuGent	64	0	1	0	1	0	0	0	1	1	1
73 0	73 0	L	SmaNuGent	78	0	0	0	0	0	0	0	0	0	1
74 0	74 0	L	SmaNuGent	78	0	1	0	1	1	1	1	0	0	0
75 0	75 0	P	SmaNuGent	74	1	1	0	0	0	0	1	1	1	1
76 0	76 1	P	SmaNuGent	74	1	1	0	0	1	1	1	0	1	0
77 0	77 0	P	SmaNuGent	76	0	1	0	0	1	0	0	1	0	1
78 1	78 0	P	SmaNuGent	76	1	1	0	0	0	1	0	0	0	1
79 0	79 0	L	SmaNuGent	74	0	0	0	0	0	0	1	1	1	1

Lampiran D: Parameter 1PL konstarin 1

Call:

```

rasch(data = DataTes()[, input$item], constraint =
cbind(ncol(DataTes()[,
input$item]) + 1, 1), IRT.param = input$TOut)

```

Model Summary:

log.Lik	AIC	BIC
-2521.99	5143.979	5264.315

Coefficients:

	value	std.errz.vals
Dffc1t.I1	-0.4661	0.2667
Dffc1t.I2	-1.2564	0.2975
Dffc1t.I3	0.1084	0.2623
Dffc1t.I4	0.9342	0.2809
Dffc1t.I5	1.4126	0.3070
Dffc1t.I6	0.0568	0.2621
Dffc1t.I7	0.5856	0.2693
Dffc1t.I8	-0.0463	0.2621
Dffc1t.I9	0.4226	0.2658
Dffc1t.I10	-1.1861	0.2935
Dffc1t.I11	0.0568	0.2621
Dffc1t.I12	1.4126	0.3070
Dffc1t.I13	-0.9878	0.2835
Dffc1t.I14	-0.4662	0.2667
Dffc1t.I15	0.1602	0.2626
Dffc1t.I16	-0.0979	0.2623
Dffc1t.I17	-0.4662	0.2667
Dffc1t.I18	0.0568	0.2621
Dffc1t.I19	1.6602	0.3259
Dffc1t.I20	-0.4124	0.2657
Dffc1t.I21	-0.6882	0.2723
Dffc1t.I22	-1.4839	0.3126

Dffclt.I23	1.5737	0.3189	4.9352
Dffclt.I24	1.1269	0.2899	3.8871
Dffclt.I25	-0.1496	0.2625	-0.5700
Dffclt.I26	3.0783	0.5320	5.7866
Dffclt.I27	-1.1860	0.2935	-4.0411
Dffclt.I28	-1.0521	0.2865	-3.6717
Dffclt.I29	-0.9250	0.2808	-3.2941
Dffclt.I30	-0.0979	0.2623	-0.3733
Dffclt.I31	-0.6880	0.2723	-2.5266
Dffclt.I32	0.1085	0.2623	0.4135
Dffclt.I33	-0.5756	0.2692	-2.1382
Dffclt.I34	3.3821	0.6053	5.5880
Dffclt.I35	1.3372	0.3021	4.4269
Dffclt.I36	0.2642	0.2635	1.0027
Dffclt.I37	-0.5756	0.2692	-2.1382
Dffclt.I38	-1.0520	0.2865	-3.6715
Dffclt.I39	-0.2016	0.2629	-0.7667
Dffclt.I40	1.7516	0.3339	5.2454
Dffclt.I41	0.1084	0.2623	0.4135
Dffclt.I42	1.4911	0.3126	4.7700
Dffclt.I43	0.1084	0.2623	0.4135
Dffclt.I44	-0.1496	0.2625	-0.5699
Dffclt.I45	0.5856	0.2693	2.1748
Dffclt.I46	-0.1496	0.2625	-0.5700
Dffclt.I47	-0.3062	0.2641	-1.1597
Dffclt.I48	0.2642	0.2635	1.0027
Dffclt.I49	1.1945	0.2935	4.0692
Dffclt.I50	-0.7457	0.2741	-2.7202
Dscrmn	1.0000	NA	NA

Integration:

method: Gauss-Hermite

quadrature points: 21

Optimization:

Convergence: 0

max(|grad|): 0.004

quasi-Newton: BFGS

Parameter 1PL konstarin 1.702

all:

```
rasch(data = DataTes()[, input$item], constraint =
cbind(ncol(DataTes()[,
input$item]) + 1, 1.702), IRT.param = input$TOut)
```

Model Summary:

log.Lik	AIC	BIC
-2559.08	5218.159	5338.495

Coefficients:

	value	std.errz.vals	
Dffc1t.I1	-0.3606	0.1413	-2.5529
Dffc1t.I2	-0.8270	0.1622	-5.1003
Dffc1t.I3	-0.0235	0.1376	-0.1708
Dffc1t.I4	0.4587	0.1493	3.0731
Dffc1t.I5	0.7379	0.1662	4.4397
Dffc1t.I6	-0.0537	0.1375	-0.3907
Dffc1t.I7	0.2552	0.1418	1.8001
Dffc1t.I8	-0.1140	0.1376	-0.8287
Dffc1t.I9	0.1601	0.1396	1.1469
Dffc1t.I10	-0.7854	0.1595	-4.9243
Dffc1t.I11	-0.0536	0.1375	-0.3900
Dffc1t.I12	0.7379	0.1662	4.4397
Dffc1t.I13	-0.6683	0.1529	-4.3708
Dffc1t.I14	-0.3605	0.1413	-2.5525
Dffc1t.I15	0.0068	0.1377	0.0491

Dffclt.I16	-0.1443	0.1378	-1.0472
Dffclt.I17	-0.3606	0.1413	-2.5527
Dffclt.I18	-0.0537	0.1375	-0.3908
Dffclt.I19	0.8828	0.1783	4.9509
Dffclt.I20	-0.3290	0.1405	-2.3412
Dffclt.I21	-0.4912	0.1452	-3.3828
Dffclt.I22	-0.9617	0.1719	-5.5942
Dffclt.I23	0.8321	0.1738	4.7877
Dffclt.I24	0.5709	0.1551	3.6803
Dffclt.I25	-0.1746	0.1380	-1.2651
Dffclt.I26	1.7142	0.3051	5.6192
Dffclt.I27	-0.7854	0.1595	-4.9243
Dffclt.I28	-0.7062	0.1549	-4.5592
Dffclt.I29	-0.6313	0.1511	-4.1787
Dffclt.I30	-0.1443	0.1378	-1.0472
Dffclt.I31	-0.4912	0.1452	-3.3826
Dffclt.I32	-0.0235	0.1376	-0.1709
Dffclt.I33	-0.4250	0.1430	-2.9715
Dffclt.I34	1.8928	0.3491	5.4215
Dffclt.I35	0.6937	0.1630	4.2566
Dffclt.I36	0.0675	0.1382	0.4888
Dffclt.I37	-0.4250	0.1430	-2.9715
Dffclt.I38	-0.7062	0.1549	-4.5593
Dffclt.I39	-0.2051	0.1384	-1.4820
Dffclt.I40	0.9363	0.1834	5.1052
Dffclt.I41	-0.0235	0.1376	-0.1708
Dffclt.I42	0.7838	0.1698	4.6167
Dffclt.I43	-0.0235	0.1376	-0.1707
Dffclt.I44	-0.1746	0.1380	-1.2649
Dffclt.I45	0.2552	0.1418	1.8000
Dffclt.I46	-0.1746	0.1380	-1.2650
Dffclt.I47	-0.2665	0.1393	-1.9135
Dffclt.I48	0.0676	0.1382	0.4891

```
Dffc1t.I49  0.6105  0.1575  3.8766
Dffc1t.I50 -0.5252  0.1465 -3.5853
Dscrmn      1.7020      NA      NA
```

Integration:

```
method: Gauss-Hermite
quadrature points: 21
```

Optimization:

```
Convergence: 0
max(|grad|): 0.0024
quasi-Newton: BFGS
```

Parameter 1PL tanpa konstrain

Call:

```
rasch(data = DataTes()[, input$item], IRT.param = input$TOut)
```

Model Summary:

log.Lik	AIC	BIC
-2482.319	5066.638	5189.38

Coefficients:

	value	std.errz	vals
Dffc1t.I1	-1.8703	0.9151	-2.0438
Dffc1t.I2	-5.1361	0.2594	-19.7989
Dffc1t.I3	0.5792	0.9720	0.5959
Dffc1t.I4	3.9618	0.6348	6.2414
Dffc1t.I5	6.0635	NaN	NaN
Dffc1t.I6	0.4523	0.9745	0.4642
Dffc1t.I7	2.6934	0.8346	3.2271
Dffc1t.I8	-0.0662	0.9786	-0.0677
Dffc1t.I9	2.0024	0.9010	2.2225
Dffc1t.I10	-4.8660	0.3940	-12.3497

Dffclt.I11	0.2406	0.9774	0.2461
Dffclt.I12	6.0782	NaN	NaN
Dffclt.I13	-4.0829	0.6223	-6.5604
Dffclt.I14	-1.7822	0.9212	-1.9346
Dffclt.I15	0.7829	0.9667	0.8099
Dffclt.I16	-0.3228	0.9771	-0.3303
Dffclt.I17	-1.8470	0.9167	-2.0148
Dffclt.I18	0.5206	0.9732	0.5350
Dffclt.I19	7.9313	NaN	NaN
Dffclt.I20	-1.6521	0.9296	-1.7773
Dffclt.I21	-2.7407	0.8358	-3.2791
Dffclt.I22	-6.2279	NaN	NaN
Dffclt.I23	6.7826	NaN	NaN
Dffclt.I24	5.1588	0.1884	27.3799
Dffclt.I25	-0.4946	0.9747	-0.5074
Dffclt.I26	13.0416	NaN	NaN
Dffclt.I27	-4.8355	0.4061	-11.9057
Dffclt.I28	-4.3361	0.5622	-7.7122
Dffclt.I29	-3.8806	0.6644	-5.8408
Dffclt.I30	-0.2529	0.9777	-0.2587
Dffclt.I31	-2.7084	0.8394	-3.2266
Dffclt.I32	0.6257	0.9709	0.6444
Dffclt.I33	-2.1691	0.8919	-2.4320
Dffclt.I34	18.3281	NaN	NaN
Dffclt.I35	5.7110	NaN	NaN
Dffclt.I36	1.6008	0.9293	1.7225
Dffclt.I37	-2.2311	0.8866	-2.5164
Dffclt.I38	-4.3754	0.5520	-7.9260
Dffclt.I39	-0.3763	0.9764	-0.3854
Dffclt.I40	7.4979	NaN	NaN
Dffclt.I41	0.5682	0.9722	0.5845
Dffclt.I42	6.6436	NaN	NaN
Dffclt.I43	0.8214	0.9655	0.8508

```
Dffc1t.I44 -0.2056  0.9781  -0.2102
Dffc1t.I45  2.4766  0.8580   2.8865
Dffc1t.I46 -0.4545  0.9753  -0.4660
Dffc1t.I47 -1.1493  0.9555  -1.2028
Dffc1t.I48  1.2676  0.9477   1.3375
Dffc1t.I49  6.1056      NaNNaN
Dffc1t.I50 -2.9492  0.8110  -3.6364
Dscrmn      0.2287      NaNNaN
```

Integration:

method: Gauss-Hermite

quadrature points: 21

Optimization:

Convergence: 0

max(|grad|): 16

quasi-Newton: BFGS

Lampiran E:Summary of the fixed effects estimates :

	Estimate	Std. Error	t-value	Pr(> t)
NilaiBhsIndo	0.004652	0.006962	0.668	0.50403
variableI1	0.097263	0.579492	0.168	0.86672
variableI2	0.849768	0.597238	1.423	0.15486
variableI3	-0.446744	0.577217	-0.774	0.43900
variableI4	-1.231605	0.588067	-2.094	0.03629 *
variableI5	-1.690440	0.603966	-2.799	0.00515 **
variableI6	-0.397888	0.577104	-0.689	0.49058
variableI7	-0.899218	0.581286	-1.547	0.12195
variableI8	-0.300293	0.577063	-0.520	0.60283
variableI9	-0.744462	0.579265	-1.285	0.19880
variableI10	0.782467	0.594810	1.315	0.18842
variableI11	-0.397888	0.577104	-0.689	0.49058
variableI12	-1.690440	0.603966	-2.799	0.00515 **
variableI13	0.593086	0.588956	1.007	0.31399
variableI14	0.097263	0.579492	0.168	0.86672
variableI15	-0.495716	0.577393	-0.859	0.39064
variableI16	-0.251437	0.577136	-0.436	0.66310
variableI17	0.097263	0.579492	0.168	0.86672
variableI18	-0.397888	0.577104	-0.689	0.49058
variableI19	-1.929738	0.616028	-3.133	0.00175 **
variableI20	0.046281	0.578941	0.080	0.93629
variableI21	0.307785	0.582563	0.528	0.59730
variableI22	1.068161	0.606511	1.761	0.07829 .
variableI23	-1.845923	0.611465	-3.019	0.00255 **
variableI24	-1.415589	0.593432	-2.385	0.01710 *
variableI25	-0.202465	0.577271	-0.351	0.72581
variableI26	-3.319775	0.772472	-4.298	1.77e-05 ***
variableI27	0.782467	0.594810	1.315	0.18842
variableI28	0.654352	0.590699	1.108	0.26803
variableI29	0.533425	0.587390	0.908	0.36387
variableI30	-0.251437	0.577136	-0.436	0.66310
variableI31	0.307785	0.582563	0.528	0.59730
variableI32	-0.446744	0.577217	-0.774	0.43900
variableI33	0.201037	0.580842	0.346	0.72928
variableI34	-3.620203	0.835991	-4.330	1.52e-05 ***
variableI35	-1.617770	0.600861	-2.692	0.00712 **
variableI36	-0.594250	0.577937	-1.028	0.30390

variableI37	0.201037	0.580842	0.346	0.72928
variableI38	0.654352	0.590699	1.108	0.26803
variableI39	-0.153317	0.577470	-0.265	0.79064
variableI40	-2.018452	0.621289	-3.249	0.00117 **
variableI41	-0.446744	0.577217	-0.774	0.43900
variableI42	-1.766340	0.607476	-2.908	0.00366 **
variableI43	-0.446744	0.577217	-0.774	0.43900
variableI44	-0.202465	0.577271	-0.351	0.72581
variableI45	-0.899218	0.581286	-1.547	0.12195
variableI46	-0.202465	0.577271	-0.351	0.72581
variableI47	-0.054247	0.578066	-0.094	0.92524
variableI48	-0.594250	0.577937	-1.028	0.30390
variableI49	-1.480647	0.595637	-2.486	0.01297 *
variableI50	0.362509	0.583581	0.621	0.53452

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Note: P-values are based on 4049 degrees of freedom

Summary of the random effects estimates:

	Estimate	Std. Error
(Intercept) AsalSekolah:SmaNegriPurwo	-0.0014	0.0127
(Intercept) AsalSekolah:SmaNuGent	-0.0003	0.0126
(Intercept) AsalSekolah:SmaPgriTeg	0.0017	0.0127

DISPERSION MODEL

NOTE: h-likelihood estimates through EQL can be biased.

Dispersion parameter for the mean model:

[1] 1.22814

Model estimates for the dispersion term:

Link = log

Effects:

Estimate	Std. Error
0.2055	0.0222

Dispersion = 1 is used in Gamma model on deviances to calculate the standard error(s).

Dispersion parameter for the random effects:

[1] 9.461e-05

Dispersion model for the random effects:

Link = log

Effects:

Estimate	Std. Error
-9.2658	6.1012

Dispersion = 1 is used in Gamma model on deviances to calculate the standard error(s).

LOG-LIKELIHOODS

h-likelihood: -2013.358

Adjusted profile likelihood

 Profiled over random effects: -2024.533

 Profiled over fixed and random effects: -2047.466

Conditional AIC: 4151.054

HGLM Model 2

ata= DataBWI

Fix-Random (fmodel)= value ~0+ Gender+NilaiBhsIndo +(1| variable)

Family Respon (fmhg)= binomial

Call:

hglm2.formula(meanmodel = fmodel, data = DataC(), family = fmhg()),

```
calc.like = TRUE, verbose = FALSE)
```

```
-----
MEAN MODEL
-----
```

Summary of the fixed effects estimates:

	Estimate	Std. Error	t-value	Pr(> t)
GenderL	-0.450734	0.528149	-0.853	0.393
GenderP	-0.520666	0.533680	-0.976	0.329
NilaiBhsIndo	0.004879	0.006829	0.714	0.475

Note: P-values are based on 4052 degrees of freedom

Summary of the random effects estimates:

	Estimate	Std. Error
(Intercept) variable:I1	0.5270	0.2708
(Intercept) variable:I10	1.1421	0.2918
(Intercept) variable:I11	0.0674	0.2670
(Intercept) variable:I12	-1.1013	0.3029
(Intercept) variable:I13	0.9762	0.2843
(Intercept) variable:I14	0.5270	0.2708
(Intercept) variable:I15	-0.0238	0.2674
(Intercept) variable:I16	0.2038	0.2671
(Intercept) variable:I17	0.5270	0.2708
(Intercept) variable:I18	0.0674	0.2670
(Intercept) variable:I19	-1.2994	0.3159
(Intercept) variable:I2	1.2000	0.2948
(Intercept) variable:I20	0.4800	0.2700
(Intercept) variable:I21	0.7198	0.2754
(Intercept) variable:I22	1.3838	0.3054
(Intercept) variable:I23	-1.2311	0.3112
(Intercept) variable:I24	-0.8637	0.2901
(Intercept) variable:I25	0.2494	0.2673
(Intercept) variable:I26	-2.2016	0.4026
(Intercept) variable:I27	1.1421	0.2918
(Intercept) variable:I28	1.0303	0.2866

(Intercept) variable:I29	0.9231	0.2822
(Intercept) variable:I3	0.0219	0.2672
(Intercept) variable:I30	0.2038	0.2671
(Intercept) variable:I31	0.7198	0.2754
(Intercept) variable:I32	0.0219	0.2672
(Intercept) variable:I33	0.6223	0.2728
(Intercept) variable:I34	-2.3292	0.4184
(Intercept) variable:I35	-1.0394	0.2993
(Intercept) variable:I36	-0.1154	0.2683
(Intercept) variable:I37	0.6223	0.2728
(Intercept) variable:I38	1.0303	0.2866
(Intercept) variable:I39	0.2951	0.2676
(Intercept) variable:I4	-0.7001	0.2829
(Intercept) variable:I40	-1.3704	0.3211
(Intercept) variable:I41	0.0219	0.2672
(Intercept) variable:I42	-1.1651	0.3068
(Intercept) variable:I43	0.0219	0.2672
(Intercept) variable:I44	0.2494	0.2673
(Intercept) variable:I45	-0.3977	0.2733
(Intercept) variable:I46	0.2494	0.2673
(Intercept) variable:I47	0.3870	0.2686
(Intercept) variable:I48	-0.1154	0.2683
(Intercept) variable:I49	-0.9208	0.2929
(Intercept) variable:I5	-1.1013	0.3029
(Intercept) variable:I50	0.7695	0.2768
(Intercept) variable:I6	0.0674	0.2670
(Intercept) variable:I7	-0.3977	0.2733
(Intercept) variable:I8	0.1583	0.2670
(Intercept) variable:I9	-0.2548	0.2703

DISPERSION MODEL

NOTE: h-likelihood estimates through EQL can be biased.

Dispersion parameter for the mean model:

[1] 1.229041

Model estimates for the dispersion term:

Link = log

Effects:

Estimate	Std. Error
0.2062	0.0222

Dispersion = 1 is used in Gamma model on deviances to calculate the standard error(s).

Dispersion parameter for the random effects:

[1] 0.8145

Dispersion model for the random effects:

Link = log

Effects:

Estimate	Std. Error
-0.2052	0.2109

Dispersion = 1 is used in Gamma model on deviances to calculate the standard error(s).

LOG-LIKELIHOODS

h-likelihood: -2089.318

Adjusted profile likelihood

 Profiled over random effects: -2111.62

 Profiled over fixed and random effects: -2118.461

Conditional AIC: 4147.953

EQL estimation converged in 2 iterations.

Lampiran F : Hasil Analisis Soal

No	No Item	Tingkat Kesulitan	Daya Pembeda	Keterangan	Subjektivitas Guru
1	I1	-7,713	0,6571	Baik	Dapat Digunakan
2	I2	-6,0706	0,1997	Soal Tidak Baik	Perlu Perbaikan
3	I3	0,1113	0,7679	Baik	Dapat Digunakan
4	I4	3,1042	0,2884	Baik	Dapat Digunakan
5	I5	-3,5498	0,3918	Soal Tidak Baik	Perlu Perbaikan
6	I6	0,1364	0,3244	Baik	Dapat Digunakan
7	I7	1,7655	0,3159	Baik	Dapat Digunakan
8	I8	-4,4031	0,0111	Soal Tidak Baik	Perlu Perbaikan
9	I9	-14,4031	0,0270	Soal Tidak Baik	Perlu Perbaikan
10	I10	7,6077	0,1492	Baik	Dapat Digunakan
11	I11	0,0400	0,7627	Baik	Dapat Digunakan
12	I12	1,2989	1,3213	Baik	Dapat Digunakan
13	I13	-1,1515	1,0016	Soal Tidak Baik	Perlu Perbaikan
14	I14	0,8899	0,5201	Baik	Dapat Digunakan
15	I15	0,3203	0,4499	Baik	Dapat Digunakan
16	I16	0,2772	0,3945	Baik	Dapat Digunakan
17	I17	5,4832	0,0813	Baik	Dapat Digunakan
18	I18	0,1643	0,3461	Baik	Dapat Digunakan
19	I19	-16,2414	0,0976	Soal Tidak Baik	Perlu Perbaikan
20	I20	0,5175	0,8218	Baik	Dapat Digunakan
21	I21	-56,4530	0,0116	Soal Tidak Baik	Perlu Perbaikan
22	I22	2,5743	0,5863	Baik	Dapat Digunakan
23	I23	-6,2882	0,2415	Soal Tidak Baik	Perlu Perbaikan
24	I24	-1,5103	0,8122	Soal Tidak Baik	Perlu Perbaikan
25	I25	0,2309	0,6186	Baik	Dapat Digunakan
26	I26	-48,1484	0,0617	Soal Tidak Baik	Perlu Perbaikan
27	I27	1,2814	1,0451	Baik	Dapat Digunakan
28	I28	1,9885	0,5313	Baik	Dapat Digunakan
29	I29	-2,4012	0,3825	Soal Tidak Baik	Perlu Perbaikan
30	I30	0,2210	0,4212	Baik	Dapat Digunakan
31	I31	-6,6487	0,0992	Baik	Dapat Digunakan
32	I32	0,4506	0,2120	Baik	Dapat Digunakan
33	I33	0,7918	0,7521	Baik	Dapat Digunakan
34	I34	-13,7209	0,2405	Soal Tidak Baik	Perlu Perbaikan
35	I35	3,2199	0,4063	Baik	Dapat Digunakan
36	I36	0,3552	0,8965	Baik	Dapat Digunakan
37	I37	-2,6097	0,2144	Soal Tidak Baik	Perlu Perbaikan
38	I38	5,4611	0,1847	Soal Tidak Baik	Perlu Perbaikan

39	I39	0,1389	1,3651	Baik	Dapat Digunakan
40	I40	16,3591	0,1022	Soal Tidak Baik	Perlu Perbaikan
41	I41	0,2239	0,4165	Baik	Dapat Digunakan
42	I42	-1,0083	-2,7773	Soal Tidak Baik	Perlu Perbaikan
43	I43	0,1994	0,6120	Baik	Dapat Digunakan
44	I44	0,0237	-2,6567	Baik	Dapat Digunakan
45	I45	0,5269	1,2356	Baik	Dapat Digunakan
46	I46	0,0272	-2,5373	Baik	Dapat Digunakan
47	I47	0,2218	-1,4634	Baik	Dapat Digunakan
48	I48	2,8725	0,0851	Soal Tidak Baik	Perlu Perbaikan
49	I49	0,7357	-5,2994	Baik	Dapat Digunakan
50	I50	0,6761	-1,2881	Baik	Dapat Digunakan

