

**EVALUASI PEMBELAJARAN MENGGUNAKAN
ITEM RESPONSE THEORY MODEL 2
PARAMETER LOGISTIK**

SKRIPSI

Oleh:
BAYU ANDRIAWAN
NIM. 14650014



**JURUSAN TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2019**

HALAMAN PENGAJUAN

**EVALUASI PEMBELAJARAN MENGGUNAKAN
ITEM RESPONSE THEORY MODEL 2
PARAMETER LOGISTIK**

SKRIPSI

**Diajukan Kepada:
Fakultas Sains dan Teknologi
Universitas Islam Negeri (UIN)
Maulana Malik Ibrahim Malang
Untuk Memenuhi Salah Satu Persyaratan Dalam
Memperoleh Gelar Sarjana Komputer (S.Kom)**

**Oleh :
BAYU ANDRIAWAN
NIM. 14650014**

**JURUSAN TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2019**

HALAMAN PERSETUJUAN

**EVALUASI PEMBELAJARAN MENGGUNAKAN
ITEM RESPONSE THEORY MODEL 2
PARAMETER LOGISTIK**

SKRIPSI

Oleh:
BAYU ANDRIAWAN
NIM. 14650014

Telah diperiksa dan disetujui untuk diuji :

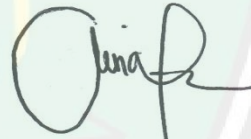
Tanggal 14 Januari 2019

Pembimbing I



Fatchurrochman, M.Kom
NIP. 19700731200501 1 002

Pembimbing II



Ainatul Mardhiyah, M.SC
NIDT. 1986033020160801 2 075

Mengetahui,
Ketua Jurusan Teknik Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang



Dr. Cahyo Crysdiyan, M.CS
NIP. 19740424 200901 1 008

HALAMAN PENGESAHAN

**EVALUASI PEMBELAJARAN MENGGUNAKAN
ITEM RESPONSE THEORY MODEL 2
PARAMETER LOGISTIK**

SKRIPSI

Oleh:
BAYU ANDRIAWAN
NIM. 14650014

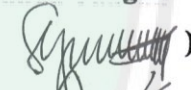
Telah Dipertahankan di Depan Dewan Penguji Skripsi dan
Dinyatakan Diterima Sebagai Salah Satu Persyaratan Untuk
Memperoleh Gelar Sarjana Komputer (S.Kom)


Tanggal 14 Januari 2019

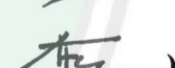
Susunan Dewan Penguji

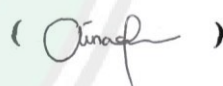
1. Penguji Utama : A'la Syaugi, M.Kom
NIP. 1977121200801 1007
2. Ketua Penguji : Supriyono, M.Kom
NIDT. 19841010 20160801 1 078
3. Sekretaris Penguji : Fathurrochman, M.Kom
NIP. 19700731 200501 1 002
4. Anggota Penguji : Ainatul Mardhiah, M.CS
NIP. 19860330 20160801 2 075

Tanda Tangan

()

()

()

()

Mengetahui,
Ketua Jurusan Teknik Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang




Dr. Agus Crysdiyan
NIP. 19740424 200901 1 008

HALAMAN PERNYATAAN

Saya yang bertanda tangan dibawah ini :

Nama : BAYU ANDRIAWAN

NIM : 14650014

Jurusan : Teknik Informatika

Fakultas : Sains dan Teknologi

Judul : **EVALUASI PEMBELAJARAN MENGGUNAKAN *ITEM RESPONSE THEORY* MODEL 2 PARAMETER LOGISTIK**

Menyatakan dengan sebenarnya bahwa skripsi yang saya tulis ini benar-nenar merupakan hasil karya saya sendiri, bukan merupakan pengambil alihan data, tulisan atau pikiran orang lain yang saya akui sebagai hasil tulisan atau pikiran saya sendiri, kecuali dengan mencantumkan sumber cuplikan pada daftar pustaka.

Apabila dikemudian hari terbukti atau dapat dibuktikan skripsi ini hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang, 26 Desember 2018
Yang Membuat Pernyataan



Bayu Andriawan
NIM.14650014

HALAMAN PERSEMBAHAN

Alhamdulillah Roobbil ‘Alamin. Puji syukur kehadiran Allah SWT yang memberikan kekuatan kepada penulis sehingga dapat menyelesaikan Kuliah S1 di kampus (Ulul Albab) Universitas Islam Negeri Maulana Malik Ibrahim Malang ini. Sholawat serta salam semoga Allah limpahkan kepada Nabi Muhammad SAW, kepada Sahabat, Keluarga, Tabiin, yang mana kita nantikan barokah syafaatnya di hari kiamat kelak.

Terimakasih kepada kedua orang tua saya, Ayah sekaligus guru saya, Bapak **SUBANDI** yang dengan kesabarannya telah mendidik saya dari kecil hingga sekarang bisa menyelesaikan pendidikan di Perguruan Tinggi dan tiada lelah bekerja mencari rizqi untuk keluarga, rizki yang berkah, rizki yang Allah ridhoi. Untuk Ibu saya tercinta **ISTIQOMAH** yang setiap hari mendoakan saya, mengiringi setiap langkah saya, mendidik saya dengan penuh kasih sayang dari lahir hingga sekarang sehingga saya mampu menyelesaikan segala kewajiban di bangku pendidikan. Semoga Allah senantiasa berikan kesehatan, panjangkan umur keduanya, dan berikan kesempatan kepada saya bahagiakan keduanya. *Aamiin.*

KATA PENGANTAR

Bismillahirrohmanirrohim

Alhamdulillah puji syukur kehadiran Allah SWT yang telah memberikan Rahmat dan Nikmat, nikmat islam, iman, ihsan, serta kekuatan sehingga penulis dapat menyelesaikan skripsi dengan baik. Shalawat serta salam semoga Allah limpahkan kepada nabi Muhammad SAW, yang kita nantikan syafaatnya kelak *fi yaumul qiyamah*.

Dengan segala kerendahan hati, penulis menyadari bahwa dalam penyelesaian Skripsi ini tidak lepas dari peran berbagai pihak yang telah banyak memberi bantuan, bimbingan dan dukungan. Dalam kesempatan ini, penulis menyampaikan terima kasih yang tak terhingga kepada:

1. Bapak Prof. Dr. Abdul Haris, M.Ag, selaku Rektor Universitas Islam Negeri Maulana Malik Ibrahim Malang.
2. Ibu Dr. Sri Harini, M.Si, selaku Dekan Fakultas Sains dan Teknologi.
3. Bapak Dr. Cahyo Crysdiyan, selaku Ketua Jurusan Teknik Informatika.
4. Bapak Fathurrochman, M.Kom, selaku dosen pembimbing I yang telah meluangkan waktu untuk membimbing, mengarahkan dan memberi masukan kepada penulis dalam pengerjaan skripsi ini hingga akhir.
5. Ibu Ainatul Mardhiyah, M.SC, selaku dosen pembimbing II yang senantiasa memberi masukan dan nasihat serta petunjuk dalam penyusunan skripsi ini.
6. Ibu Ririen Kusumawati, S.Si, M.Kom selaku dosen wali selama penulis menuntut ilmu di Universitas Islam Negeri Maulana Malik Ibrahim Malang
7. Seluruh dosen Teknik Informatika yang telah memberikan ilmu kepada penulis selama kuliah.
8. KH. M.Baidlowi Muschlih, selaku Pengasuh Pondok Pesantren Anwarul Huda, yang selalu penulis harapkan barokah doanya, beserta seluruh dewan pengasuh dan dewan pengurus pondok pesantren.
9. Bapak, Ibu dan keluarga tercinta yang senantiasa mendoakan penulis selama masa kuliah.
10. Teman-teman seperjuangan Teknik Informatika *Biner* 2014.

11. Sahabat saya Iqbal Fauzi, Ahmad Fathulloh, Lia Andini, Siska Puspita Sari, M.Burhanudin dan pesilat Arif, yang telah memfasilitasi dan berkontribusi atas laporan ini.
12. Teman-teman pondok pesantren Anwarul Huda, Karangbesuki-Sukun Malang.

Penulis menyadari bahwa dalam penyusunan skripsi ini masih terdapat kekurangan dan penulis berharap skripsi ini dapat memberikan manfaat kepada para pembaca dan khususnya penulis sendiri. *Aamiin*



MOTTO

إِذَا الْفَتَى حَسَبَ اعْتِقَادِهِ رُفِعَ
وَكُلُّ مَنْ لَمْ يَعْتَقِدْ لَمْ يَنْتَفِعْ

“Pemuda di ukur dari tekadnya, barangsiapa tidak mempunyai keyakinan (tekad) kuat maka tidak akan meraih (keberhasilan)”

~ Nadhom ‘Imrithi

“Bila Kau tak Tahan lelahnya Belajar, Maka kau harus menahan Perihnya Kebodohan.”

~ Imam Syafi’i

DAFTAR ISI

HALAMAN PENGAJUAN.....	ii
HALAMAN PERSETUJUAN.....	Error! Bookmark not defined.
HALAMAN PENGESAHAN.....	iv
HALAMAN PERNYATAAN	Error! Bookmark not defined.
HALAMAN PERSEMBAHAN	vi
KATA PENGANTAR	vii
MOTTO	ix
DAFTAR ISI.....	x
ABSTRAK.....	xiv
ABSTRACT.....	xv
المخلص.....	xvi
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang Masalah.....	1
1.2. Pernyataan Masalah.....	4
1.3. Tujuan Penelitian.....	4
1.4. Batasan Masalah.....	4
1.5. Manfaat Penelitian.....	4
BAB II STUDI PUSTAKA.....	6
2.1. Penelitian Terdahulu.....	6
2.2. Pengertian Evaluasi Pembelajaran	8
2.3. Pengertian <i>BEESMART</i>	9
2.4. <i>Item Respon Theory</i> (IRT).....	10
2.4.1 Tingkat kesulitan (b)	11
2.4.2 Kalibrasi dan estimasi parameter	14
2.4.3 Macam-macam komputasi dalam <i>Item Respon Theory</i> (IRT)	14
2.4.4 Model <i>logistic</i> 2 parameter	16
2.4.5 <i>Item Characteristic Curve</i> (ICC)	17
2.4.6 Estimasi tingkat kesulitan	18
2.4.7 <i>Item Information Function</i> (IIF)	20
BAB III DESAIN DAN IMPLEMENTASI	21
3.1 Desain Penelitian	21

3.3	Pengumpulan data	24
3.4	Desain Alur Sistem.....	25
3.4.1	Flowchart Untuk Menghitung Nilai Tingkat Kesulitan	25
3.4.2	Flowchart Daya Beda Butir Soal	26
3.4.3	Desain Flowchart ICC.....	27
3.4.4	Desain Flowchart <i>Item Information Function</i>	28
3.4.5	Perancangan <i>Interface</i>	28
3.6	Perancangan perhitungan <i>Item Response Theory</i> model 2 parameter <i>logistic</i>	29
3.6.1	<i>Item</i> parameter tingkat kesulitan butir soal	30
3.6.2	Daya pembeda butir soal.....	32
3.7	Perhitungan ICC model 2 parameter logistik	34
3.8	Perhitungan estimasi tingkat kesulitan butir soal	38
BAB IV UJI COBA DAN PEMBAHASAN		42
4.1	Data Soal Uji coba.....	42
4.2	Uji Coba Penentuan Tingkat Kesulitan Butir Soal Secara Otomatis dan Daya Beda Butir Soal	44
4.2.1	Uji Coba Penentuan Tingkat Kesulitan Butir Soal Secara Otomatis 44	
4.2.2	Uji Coba Daya Beda Butir Soal	47
4.3	Uji Coba Analisis Butir Soal Menggunakan IRT Model 2 Parameter Logistic.	50
4.3.1	Uji Coba <i>Item Characteristic Curve</i> (ICC).....	50
4.3.2	Uji Coba <i>Standart Error</i> (Se).....	55
4.3.3	Uji Coba <i>Item Information Function</i> (Iif).	56
4.3.4	Uji Coba Tingkat Kesulitan, Daya beda, ICC dan IIF.	58
4.4	Integrasi Evaluasi Pembelajaran Dalam Islam	69
BAB V KESIMPULAN DAN SARAN.....		73
5.1	Kesimpulan.....	73
5.2	Saran	74
DAFTAR PUSTAKA		75

DAFTAR GAMBAR

Gambar 3. 1 Desain alur sistem	21
Gambar 3. 2 Jawaban Mahasiswa	24
Gambar 3. 3 Jawaban Mahasiswa	24
Gambar 3. 4 Flowchart Menghitung Kesulitan.....	25
Gambar 3. 5 Flowchart Daya Beda Butir Soal.....	26
Gambar 3. 6 Flowchart ICC	27
Gambar 3. 7 Flowchart Item Information Function	28
Gambar 3. 8 Desain Interface Analisis Butir Soal	29
Gambar 3. 9 Item Characteristic Curve Butir Soal No 1	37
Gambar 4. 1 Data Soal Ujian	43
Gambar 4. 2 Hasil Pengerjaan Butir Soal	44
Gambar 4. 3 Informasi Dashbord.....	44
Gambar 4. 4 Kurva ICC Model 2 Parameter Logistic Pada Butir Soal No 3	52
Gambar 4. 5 Kurva ICC Model 2 Parameter Logistic Pada Butir Soal No 1 Sd 20	53
Gambar 4. 6 Analisis Butir Untuk Iterasi.....	55
Gambar 4. 7 Hasil Literasi Kemampuan Siswa Terhadap Butir Soal.....	56
Gambar 4. 8 Gambar Item Information Function	57

DAFTAR TABEL

Tabel 3. 1 Sebaran Soal dan Jawaban Siswa	22
Tabel 3. 2 Nilai tingkat kesulitan butir soal	31
Tabel 3. 3 Nilai Daya Beda Butir Soal.....	33
Tabel 3. 4 Perhitungan Logistik Model 2 Parameter Butir Soal 1 s/d 20	35
Tabel 3. 8 Perhitungan Nilai Perhitungan Probabilitas Pada Nilai Khas	37
Tabel 3. 9 Iterasi Butir Soal Ke 1	39
Tabel 3. 10 Perhitungan Item Information Function.....	40
Tabel 4. 1 Analisis Butir Untuk Tingkat Kesulitan	45
Tabel 4. 2 Perbandingan data sistem dan wawancara dosen.....	46
Tabel 4. 3 Perbandingan Data Sistem 1 dan Papertest.....	47
Tabel 4. 4 Analisis Butir Untuk Daya Beda.....	48
Tabel 4. 5 Perhitungan Model 2 Parameter Logistic Dengan Tingkatan Level -3 Sampai Dengan Level 3 Untuk Butir Soal No 3	51
Tabel 4. 6 Daya Beda Butir Soal No 1 s/d 20	54
Tabel 4. 7 Perhitungan IIF	57

ABSTRAK

Andriawan, Bayu. 2018. **Evaluasi Pembelajaran Menggunakan *Item Response Theory Model 2 Parameter Logistic* Pada BEESMART.** Skripsi Jurusan Teknik Informatika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang.

Pembimbing: (I) Fatchurrochman, M.Kom (II) Ainatul Mardhiyah, M.CS.

Keywords: Item Response Theory(IRT), Computer Adaptive Test.

Evaluasi adalah suatu proses merencanakan, memperoleh, dan menyediakan informasi yang sangat diperlukan untuk membuat alternative-alternatif keputusan (Mehrens & Lehmann, 1978:5). Dalam evaluasi belajar terdapat 4 jenis (Norman E.Gronlund) yang dalam penelitian ini evaluasi pembelajaran difokuskan pada butir soal yang menggunakan pilihan ganda dengan metode analisis butir Butir Soal menggunakan *Item Response Theory* model 2 parameter *logistic*. Sebagai alat bantu untuk menghitung dan menganalisis butir soal dalam penelitian ini objek yang digunakan adalah BEESMART. Untuk menentukan tingkat kesulitan butir soal secara otomatis pada aplikasi BEESMART yaitu melakukan klasifikasi terhadap jawaban peserta dengan menjumlahkan butir soal berdasarkan kategori benar dan salah. Selanjutnya, jumlah kategori butir soal benar dan salah yang diketahui dihitung dengan melakukan klasifikasi terhadap jumlah jawaban benar siswa berdasarkan kemampuan siswa atas dan bawah. Kemudian jumlah jawaban butir soal peserta tes yang telah di klasifikasikan berdasarkan kemampuan atas dan bawah dibagi dengan jumlah peserta yang memiliki kemampuan atas dan bawah. Dari proses tersebut maka akan dihasilkan sebuah nilai tingkat kesulitan yang selanjutnya nilai tingkat kesulitan butir soal akan diklasifikasikan kembali berdasarkan kategori tingkat kesulitan butir soal baru. Kemudian hasil dari klasifikasi kategori tingkat kesulitan butir soal baru akan digunakan untuk memperbarui status kategori tingkat kesulitan yang lama. Selain itu, fitur lain yang dimiliki dari penerapan *Item Response Theory Model 2 Parameter Logistic* pada aplikasi BEESMART yaitu mampu memberikan sebuah informasi berupa kategori tingkat kesulitan soal yang diperoleh dari hasil tes kemampuan siswa dengan informasi kategori butir soal yaitu mudah, sedang, susah. Selain informasi berupa tingkat kesulitan soal, terdapat juga informasi berupa kategori tingkat kelayakan butir soal, dengan keterangan informasi butir soal berupa butir soal baik, butir soal cukup baik, butir soal kurang baik dan butir soal ditolak yang digambarkan dengan keterangan dan informasi dalam bentuk kurva sehingga mampu mengetahui dayabeda dan tingkat kesulitan butir soal guna untuk meningkatkan kualitas butir soal dan mendapatkan informasi butir soal secara cepat.

ABSTRACT

Andriawan, Bayu. 2018. **Evaluation of Learning Using Item Response Theory Model 2 Parameter Logistics on BEESMART.** Department of Informatics Engineering. Faculty of Science and Technology. State Islamic University of Maulana Malik Ibrahim Malang.

Adviser: (I) Fatchurrochman, M.Kom (II) Ainatul Mardhiyah, M.CS.

Keywords: Item Response Theory (IRT), Computer Adaptive Test.

Evaluation is a process of planning, acquiring, and providing an information that needed for making decision of alternatives (Mehrens & Lehmann, 1978: 5). In the evaluation of learning there are 4 types (Norman E. Gronlund) which in this research learning evaluation focused on items that use multiple choices with Item Analysis items method. The Items using the Item Response Theory 2 models parameters logistic. As an instrument to calculate and analyze items in this research the object that used is BEESMART. To determine the difficulty level of items automatically on the BEESMART by classify of participants' answers with the addition of items based on the corect and wrong categories. The next step, number of item categories there are true and false questions is calculated by classifying the number of students' correct answers based on the ability of upper and lower students. Then the total of items answers about the test participants which has been classified according to upper and lower capabilities is divided by the total of participants who have the upper and lower abilities. From this process, the value of difficulty's level are produced. then its will be classified again based on the difficulty level category of new item. Then the results of the classification of difficulty level categories on new items will be used to update the status of the old difficulty level category. In addition, another feature possessed by the application of Item Response Theory Models 2 Parameter Logistic in the BEESMART is able to provide information about difficulty level categories of questions that obtained from the results of student's ability test with information on question of item categories namely easy, medium, difficult. In addition to information about difficulty level of the question, there is also provide information about feasibility level items, with information on the items, namely the items that are good, the items are good enough, the items are not good and the items are rejected. which is described the information by curve. so its can know the difference and the difficulty level of the items, improve the quality of items, and get information about items quickly.

الملخص

أندريان ، بايو. 2018. تقييم تعلم باستخدام نظرية الاستجابة للنموذج نموذج 2 معلمات لوجيستية في BEESMART. وصف لقسم المعلوماتية ، كلية العلوم والتكنولوجيا ، جامعة مولانا مالك إبراهيم مالانج الإسلامية الحكومية ، المشرف :

المشرفة الأولى: فتح الرحمن ، الماجستير ، عنة المرضه الثاني: خدجة فهم، الماجستير.

الكلمات الرئيسية: نظرية استجابة البند (IRT) ، اختبار التكيف الكمبيوتر.

التقييم هو عملية تخطيط واكتساب وتقديم معلومات لا غنى عنها لصنع بدائل القرار (Mehrens & Lehmann, 1978: 5). في تقييم التعلم هناك 4 أنواع (Norman E. Gronlund) والتي ركزت في هذه الدراسة على تقييم التعلم على العناصر التي تستخدم خيارات متعددة مع عنصر تحليل العناصر باستخدام *Item Response Theory Models 2 Parameter Logistic*. كأداة لحساب وتحليل العناصر في هذه الدراسة ، فإن الكائن المستخدم هو BEESMART. لتحديد مستوى صعوبة عناصر العناصر تلقائيًا في تطبيق BEESMART هو تصنيف إجابات المشاركين عن طريق جمع العناصر بناءً على الفئات الصحيحة والخاطئة. علاوةً على ذلك ، يتم حساب عدد فئات عناصر الأسئلة الحقيقية والخطئة المعروفة بتصنيف عدد الإجابات الصحيحة للطلاب استنادًا إلى قدرة الطلاب صعوداً ونزولاً. ثم يقسم عدد الإجابات على أسئلة حول المشاركين في الاختبار التي تم تصنيفها وفقاً للقدرة العليا والسفلى من قبل المشاركين الذين لديهم القدرات العليا والدنيا. من هذه العملية ، سيتم إنتاج مستوى من الصعوبة التي سيتم عندها إعادة تصنيف مستوى الصعوبة للبنود بناءً على فئة مستوى الصعوبة للبند الجديد. ثم يتم استخدام نتائج تصنيف فئات مستوى الصعوبة للعناصر الجديدة لتحديث حالة فئة مستوى الصعوبة القديم. بالإضافة إلى ذلك ، هناك ميزة أخرى يمتلكها تطبيق *Item Response Theory Models 2 Parameter Logistic* على تطبيق BEESMART وهي قدرة على توفير المعلومات في شكل فئة مستوى مستوى المشكلة التي تم الحصول عليها من نتائج اختبارات قدرة الطلاب مع معلومات حول فئات العناصر ، وهي سهلة ومتوسطة وصعبة. بالإضافة إلى المعلومات في شكل مستوى الصعوبة ، توجد أيضًا معلومات في شكل بنود تتعلق بالجودة ، مع معلومات حول البنود في شكل عناصر جيدة ، والبنود جيدة بما فيه الكفاية ، والبنود ليست جيدة وتم وصف البنود المرفوضة من حيث المعلومات بحيث قادرة على معرفة الفرق ومستوى صعوبة العناصر من أجل تحسين نوعية العناصر والحصول على معلومات حول العناصر بسرعة.

BAB I

PENDAHULUAN

Pada bab pendahuluan ini membahas latar belakang permasalahan, pernyataan masalah, tujuan penelitian, batasan masalah, dan manfaat penelitian.

1.1. Latar Belakang Masalah

Belajar merupakan sebuah hal yang sangat penting dan diutamakan bagi setiap umat Islam, hal ini dikarenakan dengan belajar akan mampu membedakan sebuah perkara yang halal dan haram, bukan hanya itu, dengan belajar juga dapat memotivasi manusia untuk terus melakukan sebuah riset dan inovasi-inovasi baru dalam kehidupan yang lebih baik dan *modern*. Sehingga belajar merupakan sebuah hal yang sangat penting dan dianjurkan oleh nabi Muhammad. S.A.W, sebagaimana hadist yang diriwayatkan oleh (HR.Ibnu Majah).

طَلَبُ الْعِلْمِ فَرِيضَةٌ عَلَى كُلِّ مُسْلِمٍ

Artinya:

“Menuntut ilmu adalah kewajiban setiap muslim”. (HR. Ibnu Majah, Shahih).

Dalam hadist diatas menjelaskan bahwasanya menuntut ilmu merupakan sebuah kewajiban bagi setiap muslim dalam memperoleh kehidupan yang baik. Khususnya dalam keilmuan bahasa arab. Karena sebagai muslim kitab Al-Quran diturunkan dengan bahasa arab sebagaimana yang difirmankan Allah S.W.T pada Surah Az-Zumar Ayat 27-28 yang berbunyi:

وَلَقَدْ ضَرَبْنَا لِلنَّاسِ فِي هَذَا الْقُرْآنِ مِنْ كُلِّ مَثَلٍ لَعَلَّهُمْ يَتَذَكَّرُونَ . قُرْآنًا
عَرَبِيًّا غَيْرَ ذِي عِوَجٍ لَعَلَّهُمْ يَتَّقُونَ

Artinya:

“Sesungguhnya telah kami buat kan bagi manusia dalam Al-Quran ini setiap macam perumpamaan supaya mereka dapat pelajaran. (Ialah) Al-Quran dalam bahasa arab yang tidak ada kebengkokan (didalamnya) supaya mereka bertakwa.” (QS. Az-Zumar: 27-28).

Dengan adanya hadist tersebut maka dapat membuktikan bahwa belajar merupakan sebuah hal yang sangat penting khususnya bahasa arab dan diwajibkan bagi setiap manusia pada umumnya dan umat muslim pada khususnya. Untuk memperoleh kesuksesan dalam belajar evaluasi adalah suatu hal yang sangat penting dan harus dilaksanakan oleh para pelajar, yaitu evaluasi pembelajaran *modern* telah banyak menggunakan komputer sebagai medianya (Erawan, 2010), hal itu didasarkan pada kemampuan komputer dalam perhitungan waktu yang *relative* singkat namun memiliki hasil perhitungan yang dapat diandalkan. Pada model ini komputer menyediakan sejumlah soal yang akan dikerjakan oleh siswa, kemudian hasil dari jawaban siswa akan dihitung berdasarkan benar dan salahnya, sehingga dapat diketahui nilai dan kemampuan siswa. Dalam *BEESMART* model ini termasuk dalam *classical test theory*.

BEESMART adalah sebuah simulasi pola pelaksanaan ujian yang diselenggarakan dengan menggunakan komputer sebagai medianya. *BEESMART* merupakan aplikasi yang sengaja dikembangkan dengan tampilan seperti Ujian Nasional Berbasis Komputer yang dikembangkan oleh tim puspendik, dengan tujuan agar para siswa di sekolah-sekolah di Indonesia ini bisa mengetahui, familier

dan lebih mengenal medan UNBK. Dalam aplikasi *BEESMART* peserta diberi sebuah kumpulan soal yang kemudian dijawab oleh siswa. Sebelumnya *BEESMART* memiliki sebuah fitur yang mampu dijadikan sebagai pengelompokan soal yang terdiri dari mudah, sedang dan susah. Hanya saja dalam fitur tersebut dilakukan dengan cara manual, kemudian melihat sebagai kelemahan dari *system BEESMART* penelitian ini mencoba untuk merubah fitur tersebut menjadi sebuah fitur yang mampu menggambarkan sebuah kemampuan butir soal terhadap kemampuan siswa berdasarkan *parameter* yang digunakan. Dalam merubah penggolongan soal menjadi otomatis, dibutuhkan sebuah perhitungan komputasi, dalam penelitian ini komputasi yang digunakan adalah *item response theory* atau sering dikenal dengan *item respon butir*, kemudian hasil perhitungan dari komputasi tersebut diharapkan berguna dalam membantu evaluasi belajar siswa.

Seperti yang telah diketahui, bahwasanya kampus UIN Maulana Malik Ibrahim Malang juga dikenal sebagai kampus yang mengedepankan integrasi antara *sains* dan *islam*. Selain itu UIN Maulana Malik Ibrahim Malang juga mengupayakan terwujudnya kampus dua bahasa yaitu bahasa arab dan bahasa inggris. Untuk mencapai tujuan tersebut diterapkanlah Program Pengembangan Bahasa Arab (PPBA). Sebagai bentuk dukungan mahasiswa jurusan Teknik Informatika terhadap kebijakan tersebut maka dalam penelitian ini disusun Evaluasi Pembelajaran Bahasa Arab Menggunakan *Item Response Theory Model 2 Parameter Logistik* Pada *BEESMART*. Diharapkan dari penelitian ini mampu memberikan sebuah gambaran *alternative* terhadap evaluasi pembelajaran bahasa arab di kampus UIN Maulana Malik Ibrahim Malang.

1.2. Pernyataan Masalah

Berdasarkan uraian diatas pernyataan masalah dalam penelitian ini adalah bagaimana menentukan tingkat kesulitan soal secara otomatis pada BEESMART dan menambahkan fitur analisis butir soal menggunakan *Item Response Theory* model 2 parameter logistic ?

1.3. Tujuan Penelitian

Adapun tujuan dari penelitian ini yaitu mengetahui tingkat kesulitan secara otomatis yang disertai dengan melakukan analisis butir soal dengan menggunakan *Item Response Theory* model 2 Parameter *Logistic* pada aplikasi BESMART.

1.4. Batasan Masalah

Adapun batasan masalah dalam penelitian ini adalah sebagai berikut:

1. Bentuk soal ujian berupa pilihan ganda bahasa arab . Dengan materi soal yang diperoleh dari pengajar Program Khusus Perkuliahan Bahasa Arab (PKPBA) UIN Maulana Malik Ibrahim Malang.
2. Sistem dikembangkan didasarkan pada aplikasi *BEESMART*.
3. Kategori soal ujian berupa *maharat qiraa'ah* dan *maharah khitabah*.
4. Evaluasi pembelajaran dititik beratkan pada analisis butir soal
5. Uji coba dilakukan di kelas 1 Awaliyah Pondok Pesantren Anwarul Huda Kota Malang.

1.5. Manfaat Penelitian

Adapun manfaat yang diperoleh dalam penelitian ini dapat diketahui sebagai berikut:

1. Penggunaan soal diharapkan dapat meningkatkan kualitas soal
2. Mendapatkan informasi secara cepat.

3. Untuk mengetahui bobot tingkat kesulitan terhadap butir soal.
4. Untuk mengetahui daya beda butir soal terhadap kemampuan siswa.



BAB II

STUDI PUSTAKA

Studi pustaka membahas tentang beberapa teori yang mendasari dalam penyusunan penelitian ini.

2.1. Penelitian Terdahulu

Anggreyani (2009) telah melakukan penelitian dengan judul penerapan *teory* uji klasik dan *teory response* butir dalam mengevaluasi butir soal. Pada penelitian ini menggunakan studi kasus : soal ujian akhir semester tingkat persiapan bersama institut pertanian bogor matakuliah fisika tahun ajaran 2008/2009. Diterangkan juga bahwa dalam penelitian ini metode yang dipakai adalah teori uji klasik, dan *item* butir soal model 1 parameter logistik, 2 parameter logistik, dan 3 parameter logistik, dengan *output* analisis perhitungan evaluasi perhitungan butir soal pada studi kasus soal ujian akhir semester tingkat persiapan bersama institut pertanian bogor matakuliah fisika tahun ajaran 2008/2009. Jika ditinjau dari persamaan dan perbedaan antara penelitian yang dilakukan oleh Anggreyani (2009) dengan penelitian ini maka diketahui persamaan yaitu sama-sama meneliti tentang evaluasi pembelajaran yang menggunakan metode *Item Response Theory* dan metode pengumpulan data yang sama-sama menggunakan metode *questionnaires* pada evaluasi belajar. dan diketahui pula bahwa perbedaan penelitian yang dilakukan oleh Anggreyani (2009) menggunakan soal ujian akhir semester tingkat persiapan bersama institut pertanian bogor mata kuliah fisika tahun ajaran 2008/2009 dengan menggunakan teori uji klasik dan teori respon butir dengan *output* analisis perhitungan dari hasil evaluasi belajar. Maka penelitian yang sekarang dilakukan di Pusat Perkuliahan Bahasa Arab (PPBA) UIN Maulana Malik Ibrahim Malang, dengan menggunakan metode *item response theory* model 2-parameter logistik

dengan *output* berupa analisis perhitungan evaluasi belajar yang ditanamkan pada sistem *BEESMART*.

Penelitian serupa pernah dilakukan oleh Wildan (2014) dengan risetnya yang berjudul *Mobile Computer Based Test* menggunakan metode *item response theory* untuk evaluasi pembelajaran bahasa arab. Pada penelitian ini metode yang digunakan adalah *item response theory* dengan menggunakan model 1 parameter logistik atau yang sering disebut dengan istilah *rash* model. Dari hasil penelitian ini *output* yang terbentuk adalah sebuah sistem *Computer Adaptive Test* dengan metode *item response theory* model 1 parameter logistik yang berbasis *website* dan *mobile*. Jika ditinjau dari persamaan dan perbedaan antara penelitian yang dilakukan oleh Wildan(2014) dengan penelitian ini maka diketahui persamaan yaitu sama-sama meneliti tentang evaluasi pembelajaran yang menggunakan metode *item response theory* dan metode pengumpulan data yang sama-sama menggunakan metode *questionnaires* pada evaluasi belajar. dan diketahui pula bahwa perbedaan penelitian yang dilakukan oleh Wildan (2014) dengan penelitian ini adalah jika penelitian Wildan (2014) menggunakan komputasi *item response theory* menggunakan model 1 parameter logistik maka dalam penelitian sekarang menggunakan komputasi perhitungan *item response theory* model 2 parameter logistik.

Hulya Olmus,dkk (2017) melakukan penelitian dengan risetnya yang berjudul *an evaluation of the two parameters (2-Pl) IRT models through a simulation study*. Pada penelitian ini metode yang digunakan adalah *item response theory* model 2-pl atau sering disebut dengan istilah *discrimination models* dengan *output* berupa analisis perhitungan dari hasil simulasi belajar. Jika ditinjau dari persamaan dan

perbedaan antara penelitian yang dilakukan oleh Hulya Olmus,dkk (2017) dengan penelitian ini maka diketahui persamaan yaitu sama-sama meneliti tentang evaluasi pembelajaran yang menggunakan metode *item response theory* dan metode pengumpulan data yang sama-sama menggunakan metode *questionnaires* pada evaluasi belajar. Adapun perbedaan penelitian yang dilakukan oleh Hulya Olmus,dkk (2017) dengan penelitian ini adalah penelitian Hulya olmus, dkk.(2017) melakukan simulasi belajar menggunakan *item response theory* model 2 parameter *logistik* dengan *output* berupa analisis perhitungan hasil evaluasi, maka pada penelitian sekarang menggunakan *item response theory* 2 parameter *logistik* dengan *output* evaluasi pembelajaran ditanamkan pada sistem *BEESMART*.

2.2. Pengertian Evaluasi Pembelajaran

Evaluasi adalah suatu proses merencanakan, memperoleh, dan menyediakan informasi yang sangat diperlukan untuk membuat alternative-alternatif keputusan (Mehrens & Lehmann, 1978:5).

Macam-macam jenis dan fungsi evaluasi pembelajaran dapat digolongkan sebagai berikut:

1. Penilaian formatif, yakni penilaian yang dilakukan pada setiap akhir satuan pelajaran, dan fungsinya untuk memperbaiki proses belajar-mengajar atau memperbaiki program satuan pelajaran.
2. Penilaian sumatif adalah, penilaian yang dilakukan setiap caturwulan atau semester (setelah siswa menyelesaikan satu unit atau bagian dari mata pelajaran tertentu), berfungsi untuk menentukan angka atau hasil belajar siswa dalam tahap-tahap tertentu.

3. Penilaian penempatan (placement) yang berfungsi untuk menempatkan siswa dalam situasi belajar-mengajar yang tepat.
4. Penilaian diagnostic, berfungsi untuk membantu memecahkan kesulitan belajar siswa.

Dari jenis dan macam-macam evaluasi belajar seperti yang disebutkan pada point diatas pada penelitian ini akan difokuskan pada evaluasi belajar yang menggunakan item butir soal pilihan ganda sebagai data tesnya.

2.3. Pengertian *BEESMART*

BEESMART merupakan sebuah *computer based test* yang dikembangkan dengan tampilan seperti ujian nasional berbasis komputer yang sengaja dibuat seperti layaknya ujian nasional berbasis komputer yang dikembangkan oleh tim puspendik, dengan tujuan agar para siswa di sekolah-sekolah di Indonesia ini bisa mengetahui, familier dan lebih mengenal medan UNBK (cbt*BEESMART*.com:2018).

Beberapa fitur yang mensupport *sytem BEESMART* kini diantaranya adalah sebagai berikut :

1. *Support* soal *listening* dengan audio dan video
2. Desain, fungsi dan navigasi seperti UNBK
3. Ujian bisa dilaksanakan dengan HP ataupun smartphone
4. Cetak kartu ujian, daftar hadir dan berita acara
5. Sinkronisasi data siswa, bank soal juga jawaban para siswa.

Untuk menentukan tingkat kesulitan soal pada aplikasi *BEESMART* yaitu klik bank soal pada menu *dashboard* dilanjutkan dengan *bank* soal kemudian buat *bank* soal baru pada tahap ini admin diharuskan untuk mengisi identitas soal yang

berupa nama *bank* soal, mata pelajaran, tingkat sekolah, jurusan, kelas, opsi pilihan jawaban, pilihan ganda, bobot pilihan, *essay* dan bobot *essay* jika sudah berhasil melengkapi maka akan diarahkan ke halaman daftar *bank* soal. Pada tahap ini klik *button* edit dan dilanjutkan dengan klik pada *button* pilih pilihan ganda. Sehingga admin akan diarahkan pada modul tambah soal dengan identitas butir soal yang wajib dilengkapi adalah jenis soal yang merupakan keterangan no pada butir soal, tingkat kesulitan yang merupakan kategori tingkat kesulitan butir soal pada saat dimasukkan pada fitur ini tersedia 3 opsi yaitu, mudah, sedang dan susah. Fitur selanjutnya yang perlu dilengkapi dalam modul tambah soal adalah acak soal yaitu fitur yang dimiliki BEESMART untuk mengacak soal, acak jawaban yang merupakan fitur BEESMART untuk mengacak kategori jawaban, form soal dan form jawaban butir soal.

2.4. *Item Respon Theory (IRT)*

Item respon theory atau yang sering disebut sebagai teori respon butir merupakan teori pengukuran *modern* yang biasanya digunakan dalam analisis butir soal. (Anggreyani, 2009). Tujuan utama *item response theory* adalah memberikan kesamaan antara statistik soal dan estimasi kemampuan peserta didik.

Menurut (Anggreyani, 2009) model-model karakteristik butir tergantung pada bentuk matematis fungsi karakteristik butirnya dan banyaknya parameter yang dilibatkan dalam model yang digunakan. Ada tiga macam model dalam perhitungan *Item responses theory (IRT)*, yaitu dalam perhitungan satu parameter atau yang disebut dengan istilah *rash*, perhitungan dua parameter atau yang biasa dikenal dengan istilah *discrimination*, dan perhitungan dengan menggunakan tiga

parameter atau yang disebut *guessing*. Untuk menggunakan model-model tersebut maka ada beberapa asumsi model *item response theory* yang harus dipahami.

Menurut (Hambleton et al.1991) beberapa asumsi model teori response butir adalah sebagai berikut:

- a. Kinerja seorang peserta pada suatu aitem dapat diprediksi oleh seperangkat factor kemampuan.
- b. Hubungan antara kinerja peserta pada usatu item dan seperangkat kemampuan dapat digambarkan oleh suatu fungsi yang menarik secara *monotonic* yang disebut *Item Charachteristic Function* atau *Item Characteristic Curve* (ICC).

Selain memahami asumsi tersebut perlu diketahui bahwa dalam *Item response theory parameter* soal terdiri dari parameter tingkat kesulitan, faktor daya pembeda, dan faktor tebakan.

2.4.1 Tingkat kesulitan (b)

Tingkat kesulitan yang merupakan salah satu parameter penting dalam *Item Response Theory* adalah sebuah parameter yang mampu memberikan informasi terhadap peluang dalam menjawab butir soal yang biasa disajikan dalam bentuk nilai indeks. Pada umumnya nilai indeks pada tingkat kesulitan *Item response theory* dinyatakan dalam bentuk proporsi yang besarnya antara 0 – 1. Dalam memperoleh nilai indeks ini adalah dengan melakukan perhitungan pada setiap butir soal dengan persamaan sebagai berikut.

$$b \frac{BA + BB}{L} \dots\dots\dots(2.1)$$

Keterangan :

b :Indeks tingkat kesulitan yang dicari

BA :Jumlah jawaban benar siswa yang termasuk kelompok pandai (Upper group) yang menjawab benar untuk tiap soal

BB :Jumlah jawaban benar siswa yang termasuk kelompok kurang (Lower group) yang menjawab benar untuk tiap soal

L :Jumlah siswa dari kelompok pandai dan kelompok kurang (jumlah upper group dan lower group)

Dari persamaan 2.1 hasil perhitungan nilai b memiliki arti yaitu semakin besar nilai b yang diperoleh setiap butir soal, berarti semakin banyak siswa yang mengerti akan soal tersebut, dan soal yang memiliki nilai b yang tinggi maka soal tersebut dianggap mudah, begitu pula sebaliknya, jika nilai b yang diperoleh setiap butir soal kecil maka semakin sedikit siswa yang mengerti akan materi pada butir soal tersebut dan soal akan dianggap susah oleh peserta tes.

Klasifikasi tingkat kesulitan menurut Nitko (1996) adalah sebagai berikut.

- a. Jika nilai b diantara 0,00 – 0,30 maka soal tergolong susah
- b. Jika nilai b diantara 0,31 -- 0,70 maka soal tergolong sedang, dan
- c. Jika nilai b diantara 0,71 – 1,00 maka soal tergolong mudah.

Daya pembeda (a)

Daya pembeda yang biasa disimbolkan dengan notasi a merupakan sebuah parameter yang memiliki informasi tentang kemampuan butir soal dalam membedakan peserta yang menguasai materi dalam butir soal dan peserta yang

belum menguasai materi yang tersedia dalam butir soal. Pada umumnya nilai indeks yang terdapat untuk daya pembeda berkisar antara 0 sampai dengan 2 (Hambleton, dkk, 1991). Semakin tinggi nilai daya pembeda soal maka semakin baik soal tersebut. Eliminasi berdasarkan daya diskriminasi dilakukan terhadap aitem yang memiliki nilai negatif. Aitem dengan daya diskriminasi negatif menunjukkan bahwa probabilitas menjawab benar aitem tersebut justru menurun seiring meningkatnya abilitas responden (Hambleton, dkk, 1991). Ini berarti ada sesuatu yang salah dengan aitem tersebut.

Untuk memperoleh nilai daya pembeda soal dapat diperoleh dengan fungsi sebagai berikut.

$$a = \frac{2 (BA - BB)}{N} \dots\dots\dots(2.2)$$

Keterangan :

- a : Daya pembeda soal
- BA : jumlah jawaban benar pada kelompok atas
- BB : Jumlah jawaban benar pada kelompok bawah
- N : Jumlah siswa yang mengerjakan tes

Crocker dan Algina (1986), daya pembeda soal dikalkulasikan sebagai berikut:

- a. Jika daya pembeda $> 0,4$ maka butir soal termasuk dalam kategori baik
- b. Jika nilai daya pembeda berkisar antara $0,3 >$ sampai dengan $< 0,4$ maka butir soal termasuk dalam kategori cukup baik.
- c. Jika nilai daya pembeda berkisar antara $0,2 >$ sampai dengan $\neg < 0,3$ maka butir soal termasuk dalam kategori kurang baik atau perlu diperbaiki.

- d. Jika daya pembeda $< 0,2$ maka butir soal ditolak.

Parameter tingkat kesulitan, faktor daya pembeda, dan faktor tebakan (*guessing*) ini dapat digunakan sebagai parameter pembantu dalam melakukan pendekatan yang dilakukan dengan *Item response theory*.

2.4.2 Kalibrasi dan estimasi parameter

Pada mulanya semua *item* masih belum memiliki parameter. Sehingga dalam memperoleh parameter yang akan digunakan dalam sistem maka perlulah dilakukan sebuah kalibrasi soal, yaitu menentukan parameter yang terdapat pada sebuah soal. Sehingga dari sebuah kalibrasi soal tersebut akan mendapatkan sebuah parameter *difficulty*, *discrimination* dan *guessing*. Parameter-parameter yang telah diperoleh sangatlah penting untuk menggambarkan sebuah kurva yang dapat mempermudah dalam membaca sebuah informasi tentang *item* soal. Dalam melakukan pengukuran terhadap *item* soal diperlukan soal *test* yang berguna untuk bahan uji terhadap siswa, dan siswa sebagai penjawab soal, sehingga dari pengukuran ini akan diperoleh parameter sepadan atau *constraint*.

2.4.3 Macam-macam komputasi dalam *Item Respon Theory* (IRT)

Jika dilihat dalam jumlah parameternya terdapat 3 macam proses perhitungan dalam *Item respon theory* yaitu sebagai berikut:

- a. Komputasi model 1 parameter *rasch*

Dalam proses perhitungan dengan menggunakan 1 parameter sering dikenal dengan istilah *rasch* model yaitu menggunakan kesulitan soal sebagai parameter pengukuran dalam sistem

Model *Rash* dituliskan sebagai persamaan berikut:

$$P_i(\theta) = \frac{e^{(\theta-b_i)}}{1+e^{(\theta-b_i)}} \text{ dengan } i:1,2,3,\dots,n \dots\dots\dots(2.3)$$

b. Komputasi model 2 parameter *discrimination*

Sesuai dengan namanya yaitu dalam proses perhitungan dengan menggunakan 2 parameter ini bukan hanya menggunakan kesulitan sebagai parameternya tetapi juga menggunakan daya pembeda yang digunakan untuk menunjukkan kemampuan/bobot soal dalam membedakan tingkat kemampuan siswa dalam pengerjaan soal. Menurut Hambleton, Swaminathan, & Rogers (1991: 15), persamaan dari proses perhitungan dengan menggunakan 2 parameter adalah sebagai berikut.

$$P_i(\theta) = \frac{e^{Da_i}}{1+e^{Da_i(\theta-b_i)}}, \text{ Dengan } i:1,2,3,\dots,n \dots\dots\dots(2.4)$$

c. Komputasi model 3 parameter *guession*

Dalam proses perhitungan yang menggunakan 3 parameter ini terdapat sebuah parameter tambahan yang berupa tebakan yaitu sebuah parameter yang melengkapi dua parameter sebelumnya yaitu tingkat kesulitan dan daya beda soal. Parameter tebakan digunakan untuk membedakan kesungguhan siswa dalam mengerjakan sebuah tes. Persamaan dari proses perhitungan dengan menggunakan 2 parameter adalah sebagai berikut (Hambleton, & Swaminathan, 1985 : 49; Hambleton, Swaminathan, & Rogers, 1991: 17) :

Persamaan komputasi model 3 parameter dapat tulis dengan persamaan sebagai berikut:

$$P_i(\theta) = c_i + (1 - c_i) \frac{e^{Da_i(\theta-b_i)}}{1+e^{Da_i(\theta-b_i)}}, \text{ Dengan } i:1,2,3,\dots,n \dots\dots\dots(2.5)$$

Keterangan:

$P_i(\theta)$: Probabilitas peserta *test* yang memiliki kemampuan θ dipilih secara acak dapat menjawab butir I dengan benar.

θ : Tingkat kemampuan subyek (sebagai variabel bebas)

b_i : Indeks kesulitan butir ke- i

e : Bilangan natural yang nilainya mendekati 2,718

n : Banyaknya butir dalam tes

a_i : Index daya pembeda

D : Faktor Penskala yang memiliki nilai 1,7

c_i : Indeks tebakan semu butir ke- i

Hubungan peluang menjawab benar $P_i(\theta)$ dengan tingkat kemampuan peserta (θ) dapat digambarkan sebagai kurva karakteristik butir (*Item characteristic curve*).

2.4.4 Model *logistic* 2 parameter

Model logistik 2 parameter merupakan model parameter yang melibatkan parameter tingkat kesulitan *item* dan parameter daya beda butir soal. Untuk persamaan dari proses perhitungan model logistik 2 parameter adalah sebagai berikut.

$$P(\theta) = \frac{1}{1 + e^{-L}} = \frac{1}{1 + e^{-1(\theta - b)}} \dots \dots \dots (2.6)$$

Keterangan:

e : adalah konstanta 2.718

b : Tingkat kesulitan butir soal

a : Daya beda parameter butir soal

L : $a(\theta - b)$ penyimpangan *logistics*

θ : *Level* tingkat kesulitan

Tahap pertama yang dapat dilakukan dalam melakukan perhitungan logistik *function* adalah dengan cara menentukan penyimpangan logistik dengan persamaan 2.7 berikut ini.

$$L = a(\theta - b) \dots \dots \dots (2.7)$$

Tahap selanjutnya adalah mencari nilai $\exp(-L)$ dengan persamaan 2.8.

$$\exp(-L) = \text{EXP} (L) \dots \dots \dots (2.8)$$

Tahap selanjutnya adalah mencari nilai $1 + \exp(-L)$ dengan persamaan 2.9 berikut ini.

$$1 + \text{EXP} (-L) \dots \dots \dots (2.9)$$

Tahap terakhir nilai dari $p(\theta)$ dapat diperoleh sebagai persamaan 2.10.

$$P(\theta) = 1 / (1 + \text{EXP} (-L)) \dots \dots \dots (2.10)$$

2.4.5 Item Characteristic Curve (ICC)

Berbeda dengan teori klasik yang memfokuskan pada informasi pada level tes, item response theory memfokuskan pada informasi pada level item, sehingga diharapkan dapat menutupi tingkat kekurangan yang terdapat pada teori klasik, penerapan IRT didasarkan atas beberapa prostutat yaitu: (1) kinerja seorang peserta pada suatu aitem dapat diprediksi oleh seperangkat factor kemampuan.(2) hubungan antara kinerja peserta pada usatu item dan seperangkat kemampuan dapat digambarkan oleh suatu fungsi yang menarik secara *monotonic* yang disebut *Item Characteristic Function* atau *Item Characteristic Curve* (ICC) (Hambleton et

al.1991). Jadi ICC adalah penggambaran dalam bentuk kurva yang menjelaskan antara hubungan antara kemampuan dengan *response* sebuah item.

Pada tahap ini respon yang telah dibuat oleh siswa akan memberikan sebuah nilai dengan ketentuan *response* soal yang bernilai benar akan diberikan skor 1 dan *response* butir soal yang bernilai salah akan diberikan skor 0. skor tes mentah terperinci diperoleh dengan menjumlahkan skor *item*, dengan ketentuan skor tes mentah akan selalu berupa bilangan bulat dan akan berkisar dari 0 hingga N, sehingga jika peserta ujian mengambil tes lagi, dengan asumsi mereka tidak ingat bagaimana mereka sebelumnya menjawab soal, skor tes mentah yang berbeda akan diperoleh maka secara hipotesis peserta ujian dapat mengambil uji berkali-kali dan dapatkan berbagai skor tes (Frank B Banker, 2001), dengan formulas skor yang benar menurut D.N Lawley adalah sebagai berikut.

$$TS_j = \sum_{i=1}^N P_i(\theta_j) \dots \dots \dots (2.11)$$

Keterangan:

TS_j : adalah nilai sebenarnya untuk peserta ujian dengan tingkat kemampuan θ_j .
 i : menunjukkan suatu butir dan $P_i(\theta_j)$ tergantung pada model *characteristic* kurva yang digunakan.

2.4.6 Estimasi tingkat kesulitan

$$\hat{\theta}_{s+1} = \hat{\theta}_s + \frac{\sum_{i=1}^N a_i [U_i - P_i(\hat{\theta}_s)]}{\sum_{i=1}^N a_i^2 P_i(\hat{\theta}_s) Q_i(\hat{\theta}_s)} \dots \dots \dots (2.12)$$

Keterangan:

$\hat{\theta}_s$ = adalah kemampuan perkiraan peserta ujian dalam iterasi s.
 a_i = adalah parameter diskriminasi item i, $i = 1, 2, \dots, N$.
 u_i = adalah respons yang dibuat oleh peserta ujian untuk item i.

u_i = 1 untuk jawaban butir soal bernilai benar.

u_i = 0 untuk jawaban butir soal bernilai salah.

$P_i(\theta_s)$ = adalah kemungkinan respon yang benar untuk *item* i , di bawah model kurva karakteristik *item* yang diberikan, pada tingkat kemampuan θ_j dalam iterasi s

$Q_i(\theta_s^{\wedge})$ = $1 - P_i(\theta_s)$ adalah kemungkinan respons yang salah terhadap butir i , dibawah model kurva karakteristik item yang diberikan, pada level tingkat kesulitan θ_s dalam iterasi s .

Dalam proses estimasi parameter, dicari kalkulasi terhadap theta yang paling sesuai untuk peserta. Hasil theta yang paling sesuai dengan peserta akan diperoleh dengan menghitung standart error minimum terhadap pelaksanaan test yang telah dilakukan. Karena kemampuan awal peserta belum diketahui, maka pada mulanya peserta dianggap memiliki tingkat kemampuan sedang. Dalm IRT dengan jangkauan theta berkisar antara 0 dan 1 maka theta dapat diberi nilai 0,5 nilai ini akan terus berubah ketika tes berlangsung, sedangkan SEM akan semakin menurun hingga berjalannya tes yang semakin banyaknya soal yang terjawab.

Adapun Formulai SEM dapat dihitung dengan persamaan berikut:

$$SE(\theta_s) = \frac{1}{\sqrt{\sum_{i=1}^N a_i^2 P(\theta^{\wedge}) Q(\theta^{\wedge})}} \dots\dots\dots(2.13)$$

Selain melakukna perhitungan SEM untuk mencari kalkulasi terhadap theta yang paling sesui untuk peserta, dilakukan juga proses perhitungan IIF atau Item Information Function yang bertujuan untuk memperoleh informasi keseluruhan dari

tes yang telah dilaksanakan. Adapun Formula untuk menghitung IIF adalah sebagai berikut.

2.4.7 Item Information Function (IIF)

Informasi pada IIF didefinisikan oleh *fisher* sebagai definisi dari hubungan timbal balik dari keakuratan *item* soal yang telah diestimasi (Yu:2010). Pada dua parameter model *item information function* didefinisikan sebagai persamaan berikut:

$$I_i(\theta) = a_i^2 p_i(\theta) Q_i(\theta) \dots \dots \dots (2.14)$$

Keterangan:

a_i : merupakan parameter daya beda butir i

$P(\theta)$: $1 / (1 + \text{EXP}(-L))$

$Q_i(\theta)$: $1 - P_i(\theta)$

θ : merupakan *level* tingkat kemampuan

BAB III

DESAIN DAN IMPLEMENTASI

Berdasarkan uraian yang meliputi latar belakang masalah, pernyataan masalah tujuan penelitian batasan masalah manfaat penelitian dan studi pustaka pada bab 2 maka rangkaian alur desain dan implementasi adalah sebagai berikut.

3.1 Desain Penelitian



Gambar 3. 1 Desain alur sistem

Pada tahap ini desain sistem dibuat dengan tujuan untuk memberi gambaran tentang bagaimana *item response theory* dengan model 2 parameter logistik bekerja pada sistem *BEESMART*. Untuk menerapkan *item response theory* pada sistem *BEESMART* yaitu dengan menentukan nilai b selaku tingkat kesulitan dan nilai a yang merupakan parameter daya pembeda. Selanjutnya parameter a dan b yang merupakan parameter penyusun dari parameter *item response theory* diolah kembali untuk mengetahui estimasi kemampuan siswa terhadap soal. Setelah proses estimasi kemampuan selesai dilakukan, tahap selanjutnya adalah melakukan perhitungan *Item characteristic curve*. Tahap terakhir dalam perhitungan *item response theory* adalah melakukan perhitungan terhadap *item information function*.

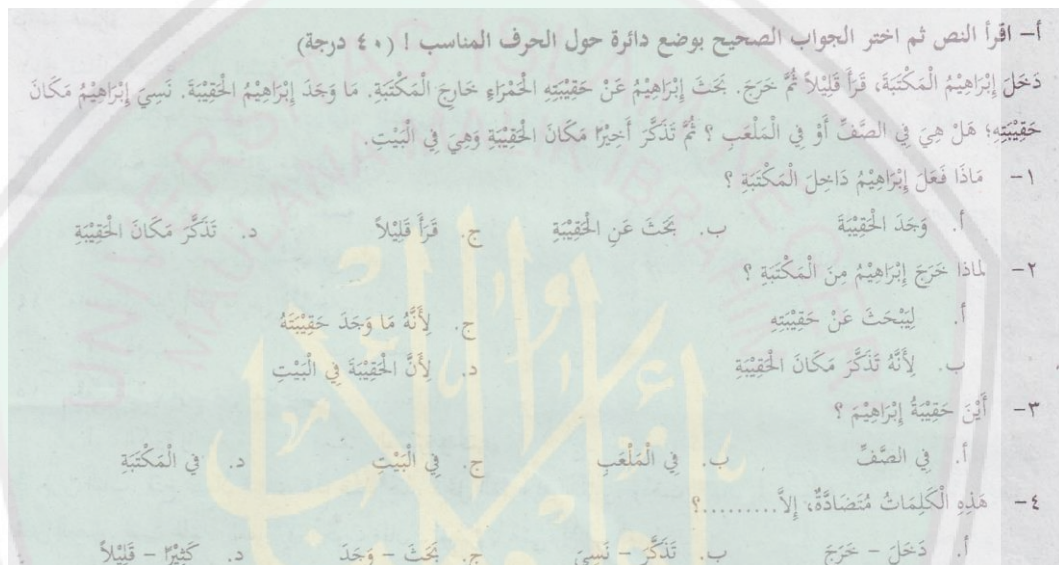
Tabel 3. 1 Sebaran Soal dan Jawaban Siswa

NO	Nama Siswa	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	BENAR	SALAH
1	Farohi Dawin N	1	1	0	0	1	1	0	0	0	1	0	0	0	0	0	0	0	1	1	0	7	13
2	Abdul Malik	1	1	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	1	1	0	6	14
3	Riyan Pratama	1	1	0	0	1	0	1	0	0	0	0	0	0	0	1	0	0	0	1	0	6	14
4	Wahyu Sadewo	1	1	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	1	1	0	6	14
5	Arya Bayu	1	1	1	0	0	0	0	0	0	0	1	0	0	0	1	0	0	1	0	0	6	14
6	Bagus F Mahendra	1	1	1	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	6	14
7	Muhammad A	0	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	1	1	0	5	15
8	Izul Azmi	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	5	15
9	Ali Murtadho	0	1	0	0	1	0	0	0	0	0	0	0	1	0	1	0	0	0	1	0	5	15
10	Sondwi Haris	0	1	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	1	1	0	5	15
11	M Farul Bari	0	1	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	1	0	5	15
12	Heri Lestari	1	1	0	0	1	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	5	15
13	Hakim Muhammad	1	1	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	4	16
14	Fajar Santosa	0	1	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	1	0	4	16
15	Illa Hidayati	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	1	0	4	16
16	Muhammad Sirojui Umam	1	1	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	4	16
17	Wildan Ansori	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	4	16
18	M Fajar	0	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	4	16
19	Abdul Fatah	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	3	17
20	Iqbal Ari Gunawan	0	1	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	3	17

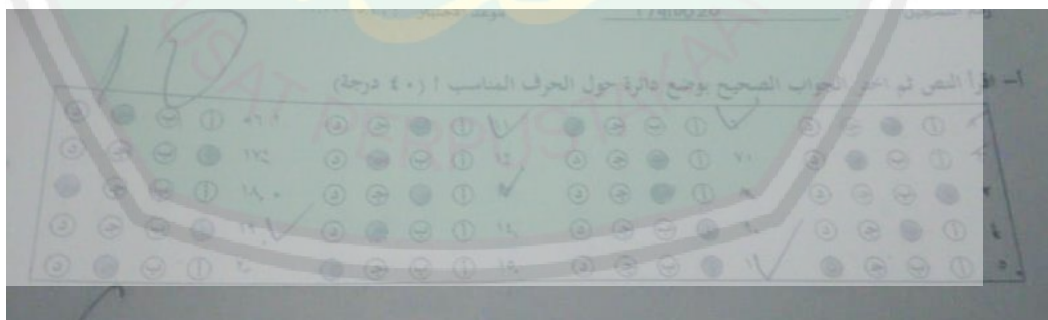
21	Albi Peranda Putra	0	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	3	17
22	Ahmad Khoirummahdun	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	17
23	Ali Umar Ritonga	0	1	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	3	17
24	Wahyu Alif	0	1	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	3	17
25	Dimas Faisal	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	3	17	
26	Alfani Hidayat	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	2	18
27	Ubaidillah	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	2	18
28	Nur Khamid	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	19
29	Rikan Istiahi	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	19

3.3 Pengumpulan data

Dalam penelitian ini, data yang digunakan adalah soal dan jawaban matakuliah bahasa arab yang peroleh dari Program perkuliahan Bahasa Arab dengan jumlah butir soal sebanyak 20, peserta ujian sebanyak 29 mahasiswa. Dengan data pembanding yaitu diambil dari seorang dosen bahasa arab yang dilakukan dengan cara wawancara. Adapaun contoh soal asli adalah dapat dilihat pada gambar berikut.



Gambar 3. 2 Jawaban Mahasiswa



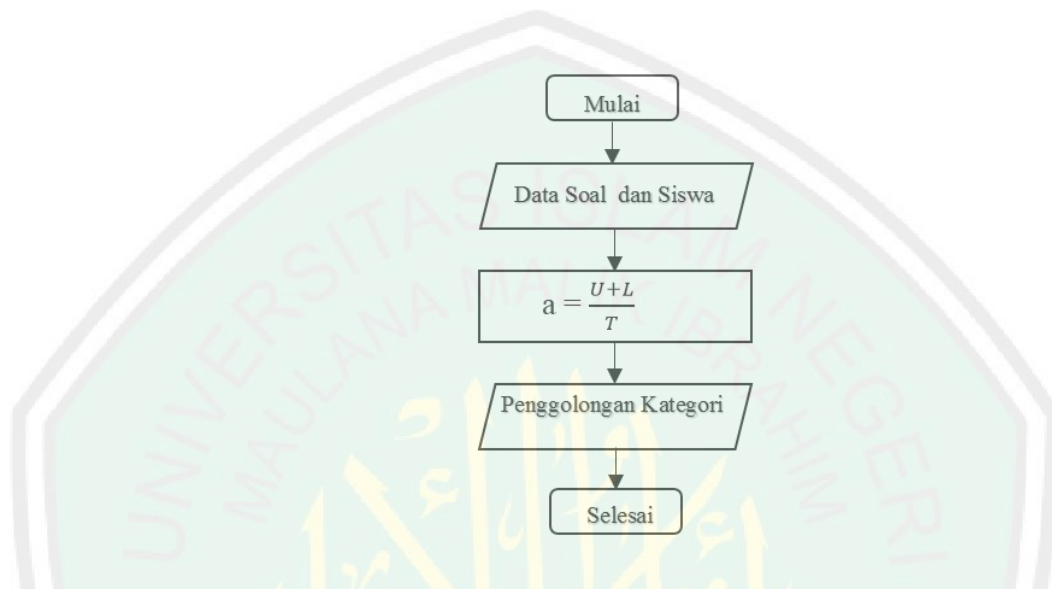
Gambar 3. 3 Jawaban Mahasiswa

Bahasa Arab Universitas Islam Negeri Maulana Malik Ibrahim Malang, dengan teknik pengumpulan data menggunakan wawancara, studi dokumen pada

kegiatan evaluasi belajar Pembelajaran Bahasa Arab Universitas Islam Negeri Maulana Malik Ibrahim Malang.

3.4 Desain Alur Sistem

3.4.1 Flowchart Untuk Menghitung Nilai Tingkat Kesulitan



Gambar 3. 4 Flowchart Menghitung Kesulitan

Tahap pertama dalam menentukan tingkat kesulitan adalah dengan menyediakan data soal yang kemudian di berikan kepada peserta tes untuk dikerjakan. Response dari jawaban yang diberikan siswa akan dihitung mengenai jumlah benar dan salahnya yang, diawali dengan membedakan kemampuan peserta atas dan bawah, jawaban benar siswa atas dijumlahkan dengan jawaban siswa kemampuan bawah yang kemudian dibagi dengan jumlah kemampuan siswa atas dan bawah. Selanjutnya nilai tingkat kesulitan yang telah diperoleh untuk semua butir akan di klasifikasikan menurut rangenya pembagiannya, adapun range pembagiannya dapat dilihat pada Nitko (1996).

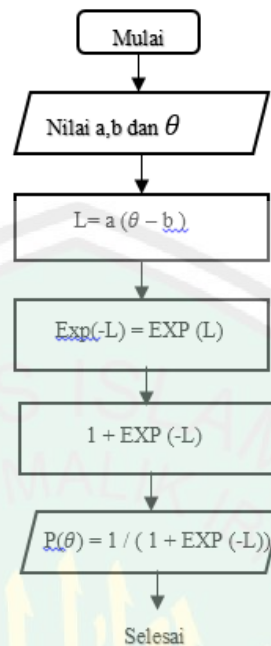
3.4.2 Flowchart Daya Beda Butir Soal



Gambar 3. 5 Flowchart Daya Beda Butir Soal

Pada desain flowchart daya beda butir soal menunjukkan nilai daya beda butir soal diperoleh dari data response jawaban siswa. Pada tahap ini jumlah nilai kelompok *upper* ditambah dengan jumlah nilai kelompok *lower* yang kemudian dibagi dengan jumlah total kelompok *upper* dan *lower*. Dari proses perhitungan ini akan diperoleh nilai daya beda untuk perbutir soal.

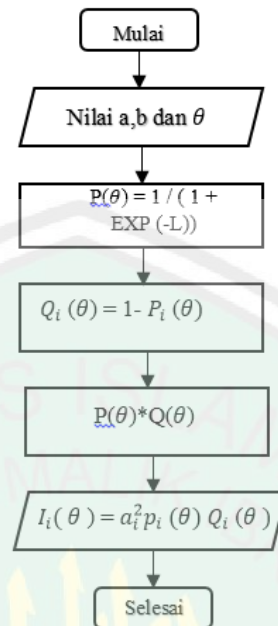
3.4.3 Desain Flowchart ICC



Gambar 3. 6 Flowchart ICC

Pada tahap ini merupakan proses awal dalam *item response theory* dengan nilai tingkat kesulitan (b) dan daya beda (a) sudah diketahui dari response siswa terhadap butir soal. Selanjutnya nilai tingkat kesulitan dan daya beda akan dihitung dengan memadukan nilai khas yang nilainya berkisar antara -3 sampai dengan 3. Sehingga dari proses perhitungan ini akan mampu membentuk sebuah grafik s yang menggambarkan kemampuan setiap butir soal terhadap siswa.

3.4.4 Desain Flowchart *Item Information Function*

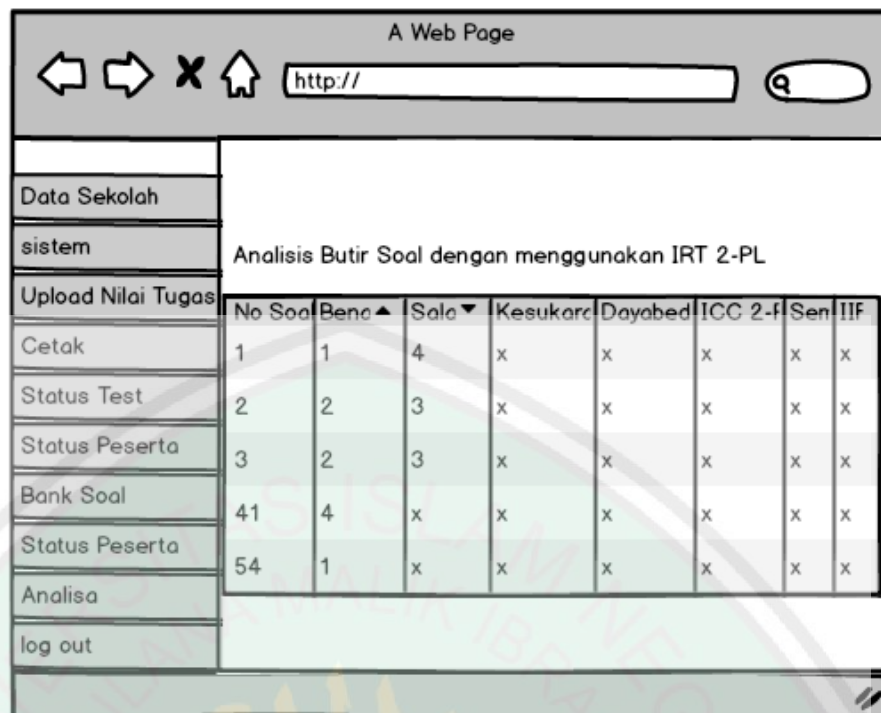


Gambar 3. 7 Flowchart Item Information Function

Pada tahap ini, nilai item information function diperoleh dari nilai a , b dan nilai khas kemampuan yang digunakan untuk mencari nilai probabilitas yang kemudian hitung guna mendapatkan nilai Q sehingga setelah diperoleh nilai P dan Q nilai item information function dapat diperoleh seperti persamaan 2.14.

3.4.5 Perancangan *Interface*

Perancangan *interface* merupakan bagian dari *software* yang digunakan oleh *enduser* dan bisa dilihat pada layar monitor apabila sebuah program dijalankan.



Gambar 3. 8 Desain Interface Analisis Butir Soal

3.6 Perancangan perhitungan *Item Response Theory* model 2 parameter logistic.

Sebuah gambaran komputasi yang akan dijalankan oleh sistem yang dapat membantu *programmer* dalam menerapkannya dalam bentuk *coding*. Untuk mempermudah pemahaman mengenai *item respon teori* model 2 parameter logistik dapat di lihat pada Tabel 3.1 tentang sebaran soal dan jawaban siswa, pada Tabel 3.1 dijelaskan bahwasannya data diperoleh dari 29 siswa dengan 20 butir soal, dengan informasi butir soal bernilai 0 jika butir soal bernilai salah, dan 1 jika butir soal bernilai benar. Pada Tabel 3.1 terdapat kolom nilai, yaitu kolom yang menggambarkan informasi jumlah butir soal benar yang dijawab soal dalam penelitian ini terdapat 20 item butir soal, sebagai contoh mahasiswa dengan nama Nailah Primasari dari butir soal 1 sampai dengan butir soal no 20, mahasiswa tersebut

berhasil menjawab butir soal sebanyak 8. Terdapat pula baris jumlah *item* benar, yaitu sebuah baris yang memiliki sebuah informasi tentang jumlah *item* butir soal yang berhasil dijawab oleh seluruh peserta tes. Dalam penelitian ini peserta tes terdapat 29 mahasiswa sebagai contoh pada butir soal no 1, yaitu dari jumlah peserta evaluasi sebanyak 29 hanya memiliki 11 soal benilai benar.

3.6.1 Item parameter tingkat kesulitan butir soal

Untuk memperoleh parameter tingkat kesulitan butir soal dapat dihitung dengan formulasi sebagai berikut.

$$b = \frac{BA+BB}{L} \dots\dots\dots(3.1)$$

Keterangan :

b :Indeks tingkat kesulitan yang dicari

BA :Jumlah siswa yang termasuk kelompok pandai (*Upper group*) yang menjawab benar untuk tiap soal

BB : Jumlah siswa yang termasuk kelompok kurang (*Lower group*) yang menjawab benar untuk tiap soal

T :Jumlah siswa dari kelompok pandai dan kelompok kurang (jumlah *upper group* dan *lower group*)

Menurut Nitko (1996) nilai tingkat kesulitan dapat diklasifikasikan sebagai berikut.

1. Jika nilai b diantara 0,00 – 0,30 maka soal tergolong susah
2. Jika nilai b diantara 0,31 -- 0,70 maka soal tergolong sedang, dan

3. Jika nilai b diantara $0,71 - 1,00$ maka soal tergolong mudah.

Misalkan dalam perhitungan ini diambil contoh perhitungan dari *item* butir soal no 1 yaitu jumlah siswa yang menjawab benar sebanyak 11 orang dengan jumlah siswa keseluruhan yaitu 29, maka berdasarkan formula *item* kesulitan yaitu jumlah butir soal benar pada kelompok atas di tambah dengan jumlah jawaban benar pada kelompok bawah di bagi dengan jumlah kelompok atas dan bawah. $(6+1) / 14 = 0,5$ tahap selanjutnya adalah melakukan klasifikasi terhadap kategori nilai, maka berdasarkan klasifikasi menurut nitko nilai item butir soal pada no 1 tergolong dalam kategori sedang, untuk lebih jelasnya dapat dilihat pada Tabel 3.2.

Tabel 3. 2 Nilai tingkat kesulitan butir soal

Butir	Tingkat Kesulitan	Keterangan
1	0.5	Sedang
2	1.071429	Mudah
3	0.142857	Susah
4	0	Susah
5	0.571429	Sedang
6	0.142857	Susah
7	0.142857	Susah
8	0	Susah
9	0	Susah
10	0.142857	Susah
11	0.071429	Susah
12	0	Susah
13	0	Susah
14	0.071429	Susah
15	0.5	Sedang
16	0	Susah
17	0	Susah

Butir	Tingkat Kesulitan	Keterangan
18	0.428571	Susah
19	0.5	Sedang
20	0	Susah

3.6.2 Daya pembeda butir soal

Daya pembeda atau *discrimination* yang biasa disimbolkan dengan notasi a merupakan sebuah parameter yang memiliki informasi tentang kemampuan butir soal dalam membedakan peserta yang menguasai butir soal dan peserta yang belum menguasai butir soal. Pada umumnya nilai indeks yang terdapat untuk daya pembeda berkisar antara -1 sampai dengan 1 semakin tinggi nilai daya pembeda soal maka semakin baik soal tersebut. Untuk memperoleh nilai daya pembeda soal dapat diperoleh dengan fungsi sebagai berikut:

$$a = \frac{2(BA - BB)}{N} \dots \dots \dots (3.2)$$

Keterangan :

a : Daya pembeda soal

BA : jumlah jawaban benar pada kelompok atas

BB :Jumlah jawaban benar pada kelompok bawah

N :Jumlah siswa yang mengerjakan tes (kelompok atas dan bawah)

Dalam pembagian kelompok golongan dilakukan dengan mengambil dari 25% urutan nilai terbaik sebagi siswa kelompok atas, dan 25% dari nilai terendah sebagai siswa kelompok bawah. Hal ini seperti yang telah dijelaskan oleh Anastasi dan Urbina (1997).

Pada Tabel 3.1 dapat diketahui bahwasanya jumlah peserta tes sebanyak 29 siswa, sehingga jika mengacu pada penelitian anastasi dan urbina (1997) maka $29 \times 25\% = 7,25$, jika dibulatkan sebanyak 7, itu berarti dari total 29 peserta *test* yang

termasuk dalam kelompok atas dan kelompok bawah masing-masing sebanyak 7 peserta. Sebagai contoh untuk butir soal no 1 peserta kelompok yang memiliki kemampuan atas memiliki 6 butir soal benar dan 1 butir soal bernilai salah sedangkan untuk peserta kelompok yang memiliki kemampuan rendah memiliki 1 butir soal bernilai benar dan 6 butir soal bernilai salah. Untuk contoh perhitungan daya beda butir soal no 1 dapat dilihat sebagai berikut.

$$a = \frac{2(BA - BB)}{N} = \frac{2(6 - 1)}{14} = 0,714$$

Sampai disini dengan cara yang sama dapat dilakukan guna untuk melihat nilai daya beda butir soal 2 sampai dengan butir soal ke 20. Untuk nilai daya beda butir soal no 1 s/d 20 dapat dilihat pada Tabel 3.3

Tabel 3. 3 Nilai Daya Beda Butir Soal

Butir Soal	Kelompok Atas	Kelompok Bawah	Daya Beda(a)
1	6	1	0.714285714
2	8	8	0
3	3	0	0.428571429
4	0	0	0
5	6	3	0.428571429
6	2	0	0.285714286
7	2	0	0.285714286
8	0	0	0
9	0	0	0
10	2	0	0.285714286
11	1	0	0.142857143
12	0	0	0
13	0	0	0
14	0	1	-0.142857143
15	4	3	0.142857143
16	0	0	0
17	0	0	0
18	6	1	0.714285714

Butir Soal	Kelompok Atas	Kelompok Bawah	Daya Beda(a)
19	7	1	0.857142857
20	0	0	0

3.7 Perhitungan ICC model 2 parameter logistik

Model logistik 2 parameter merupakan model parameter yang melibatkan parameter tingkat kesulitan *item* dan parameter daya beda butir soal. Untuk persamaan dari proses perhitungan model *logistic* 2 parameter adalah sebagai berikut.

$$P(\theta) = \frac{1}{1 + e^{-L}} = \frac{1}{1 + e^{-1(\theta - b)}} \dots \dots \dots (3.3)$$

Keterangan:

- e :Konstanta 2.718
- b :Tingkat kesulitan butir soal
- a :Daya beda parameter butir soal
- L : a(θ – b) penyimpangan *logistic*
- θ : *Level* tingkat kesulitan

Tahap pertama dalam melakukan perhitungan *logistic function* adalah dengan cara menentukan penyimpangan *logistic* dengan persamaan pada persamaan 3.4, dan untuk contoh perhitungan diambil dari butir soal no 1.

$$L = a (\theta - b) \dots \dots \dots (3.4)$$

$$L = 0.714285714 (1 - 0.5) = 0.357143$$

Tahap selanjutnya adalah mencari nilai exp(-L) dengan persamaan berikut:

$$\text{Exp}(-L) = \text{EXP} (L) \dots \dots \dots (3.5)$$

$$\text{Exp}(-L) = \text{EXP} (0.357143)$$

$$\text{Exp}(-L) = 2,718 * (0.357143) = 0,699673$$

Tahap selanjutnya adalah mencari nilai $1 + \text{exp}(-L)$ dengan persamaan berikut:

$$1 + \text{EXP} (-L) \dots\dots\dots(3.6)$$

$$1 + 0,699673 = 1,699673$$

Tahap terakhir nilai dari $p(\theta)$ dapat diperoleh sebagai persamaan berikut

$$P(\theta) = 1 / (1 + \text{EXP} (-L)) \dots\dots\dots(3.7)$$

$$P(\theta) = 1 / 1.699673 = 1.42924$$

Untuk perhitungan logic model 2 parameter *logistic* pada butir 1 sampai dengan butir 20 dapat dilihat pada Tabel 3.4 s/d 3.7

Tabel 3. 4 Perhitungan Logistik Model 2 Parameter Butir Soal 1 s/d 20

Butir Soal	Tingkat kesulitan (b)	Daya Beda (a)	Level (θ)	$L = a (\theta - b)$	$\text{exp} (-L)$	$1 + \text{exp}(-L)$	$p=1/(1+ \text{exp}(-L))$
1	0.5	0.71428571	1	0.357142857	0.699672537	1.699672537	0.588348625
2	1.071428571	-0.1428571	1	0.010204082	0.989847803	1.989847803	0.502550998
3	0.142857143	0.28571429	1	0.244897959	0.78278441	1.78278441	0.560920319
4	0	0	1	0	1	2	0.5
5	0.571428571	0.28571429	1	0.12244898	0.884751044	1.884751044	0.530574053
6	0.142857143	0	1	0	1	2	0.5
7	0.142857143	0.28571429	1	0.244897959	0.78278441	1.78278441	0.560920319
8	0	0	1	0	1	2	0.5
9	0	0	1	0	1	2	0.5
10	0.142857143	0.28571429	1	0.244897959	0.78278441	1.78278441	0.560920319
11	0.071428571	0.14285714	1	0.132653061	0.875768878	1.875768878	0.53311472
12	0	0	1	0	1	2	0.5
13	0	0	1	0	1	2	0.5
14	0.071428571	-0.1428571	1	-0.132653061	1.141853776	2.141853776	0.46688528
15	0.5	0.14285714	1	0.071428571	0.93106278	1.93106278	0.517849554
16	0	0	1	0	1	2	0.5
17	0	0	1	0	1	2	0.5
18	0.428571429	0.57142857	1	0.326530612	0.72142229	1.72142229	0.580914983
19	0.5	0.71428571	1	0.357142857	0.699672537	1.699672537	0.588348625
20	0	0	1	0	1	2	0.5

Pada tahap ini respon yang telah dibuat oleh siswa akan memberikan sebuah nilai dengan ketentuan *response* soal yang bernilai benar akan diberikan skor 1 dan *response* butir soal yang bernilai salah akan diberikan skor 0. Skor tes mentah terperinci diperoleh dengan menjumlahkan skor *item*, dengan ketentuan bahwa skor tes mentah akan selalu berupa bilangan bulat dan akan berkisar dari 0 hingga N, sehingga ketika peserta ujian mengambil tes lagi, dengan asumsi mereka tidak ingat bagaimana mereka sebelumnya menjawab soal, skor tes mentah yang berbeda akan diperoleh maka secara hipotesis, yaitu peserta ujian dapat melakukan ujian berkali-kali dan dapatkan berbagai skor tes (Frank B Banker, 2001), dengan formulas skor yang benar menurut D.N Lawley adalah sebagai berikut.

$$TS_j = \sum_{i=1}^N P_i(\theta_j) \dots \dots \dots (3.8)$$

Untuk contoh perhitungan *Item Characteristic Curve* diambil dari butir soal no 1 dengan tingkat kemampuan 1.

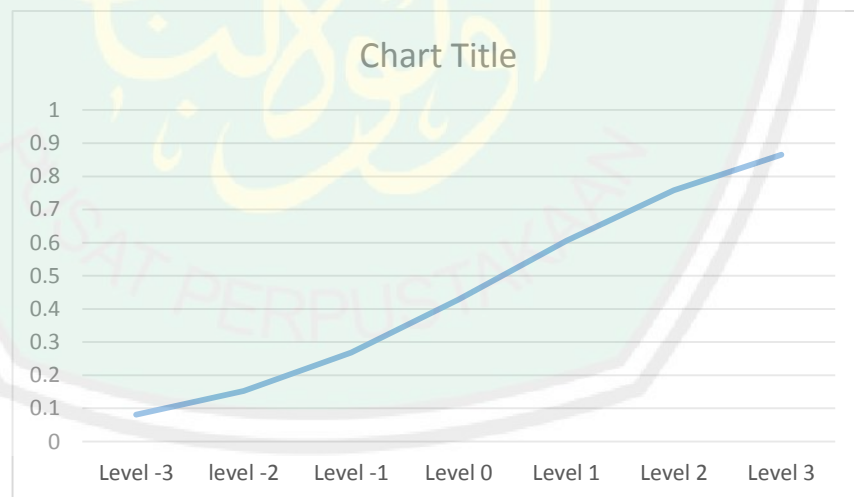
$$P = 1/(1 + \exp(-0.714285714 (1-0.5))) = 0.5883$$

Tahap selanjutnya yaitu menentukan fungsi logistic pada setiap response sebaran butir jawaban siswa dengan menggunakan parameter tingkat kesulitan yang dilambangkan dengan b untuk didefinisikan sebagai nilai pada kemampuan probabilitas response yang memiliki nilai khas yang berkisar $-3 \leq b \leq +3$ sehingga dari pengukuran dengan titik khas tersebut item characteristic curve akan terbentuk sebuah curve s yang merupakan gambaran kemampuan siswa. Contoh perhitungan nilai khas dapat dilihat pada Tabel. 3.6 dan hasil kurva perhitungan nilai khas dapat dilihat pada Gambar 3.3 *Item characteristic curve* butir soal no 1

b	a	θ	$L = a (\theta - b)$	$\exp (-L)$	$1 + \exp(-L)$	$p = 1/(1 + \exp(-L))$
0.4	0.714285714	-3	-2.42857	11.34267	12.34267	0.081019769
0.4	0.714285714	-2	-1.71429	5.552708	6.552708	0.152608665
0.4	0.714285714	-1	-1	2.718282	3.718282	0.268941421
0.4	0.714285714	0	-0.28571	1.330712	2.330712	0.429053403
0.4	0.714285714	1	0.428571	0.651439	1.651439	0.605532487
0.4	0.714285714	2	1.142857	0.318907	1.318907	0.758203828
0.4	0.714285714	3	1.857143	0.156118	1.156118	0.864963577

Tabel 3. 5 Perhitungan Nilai Perhitungan Probabilitas Pada Nilai Khas

b	a	θ	$L = a (\theta - b)$	$\exp (-L)$	$1 + \exp(-L)$	$p = 1/(1 + \exp(-L))$
0.5	0.7143	-3	-2.5	12.182	13.18249395	0.07585818
0.5	0.7143	-2	-1.7857143	5.9638	6.963838307	0.143598969
0.5	0.7143	-1	-1.0714286	2.9195	3.919547303	0.255131504
0.5	0.7143	0	-0.3571429	1.4292	2.429240032	0.411651375
0.5	0.7143	1	0.3571429	0.6997	1.699672537	0.588348625
0.5	0.7143	2	1.0714286	0.3425	1.342518855	0.744868496
0.5	0.7143	3	1.7857143	0.1677	1.167677249	0.856401031



Gambar 3. 9 Item Characteristic Curve Butir Soal No 1

Gambar 3.3 merupakan gambaran daya beda butir yang menunjukkan kemiringan kurva. Kecuraman pada lengkungan merupakan koefisien arah a pada fungsi $a(\theta - b)$. semakin curam kemiringan nilai dari kurva maka semakin besar nilai

daya beda butir soal terhadap siswa, atau dengan kata lain semakin curam kurva maka semakin mudah untuk mengetahui karakter daya beda butir soal terhadap kemampuan siswa.

3.8 Perhitungan estimasi tingkat kesulitan butir soal

Response jawaban siswa yang sudah diberikan terhadap butir soal akan dilakukan estimasi tingkat kesulitan butir soal guna memperkirakan tingkat kemampuan masing-masing siswa terhadap butir soal. Untuk persamaan dari proses perhitungan estimasi tingkat kesulitan butir soal adalah sebagai berikut.

$$\theta_{s+1}^{\wedge} = \theta_s^{\wedge} + \frac{\sum_{i=1}^N a_i [U_i - P_i(\theta_s)]}{\sum_{i=1}^N a_i^2 P_i(\theta_s^{\wedge}) Q_i(\theta_s^{\wedge})} \dots \dots \dots (3.9)$$

Keterangan:

- θ_s^{\wedge} = adalah kemampuan perkiraan peserta ujian dalam iterasi s
- a_i = adalah parameter diskriminasi item i, $i = 1, 2, \dots, N$
- u_i = adalah respons yang dibuat oleh peserta ujian untuk item i:
- u_i = 1 untuk jawaban butir soal bernilai benar
- u_i = 0 untuk jawaban butir soal bernilai salah
- $P_i(\theta_s)$ = adalah kemungkinan respon yang benar untuk item i, dibawah model kurva karakteristik item yang diberikan, pada tingkat kemampuan θ_j dalam iterasi s
- $Q_i(\theta_s^{\wedge})$ = $1 - P_i(\theta_s)$ adalah kemungkinan respons yang salah terhadap butir i, di bawah model kurva karakteristik item yang diberikan, pada level tingkat kesulitan θ_s dalam iterasi s

Untuk contoh perhitungan dapat diambil dari perkiraan kemampuan nanik sofiatun yang merupakan salahsatu peserta tes yang mengikuti tes kemampuan siswa.

Tabel 3. 6 Iterasi Butir Soal Ke 1

Item	b	a	U(respon e)	p	q	a*(u-p)	a*a*(p*q)
1	0.5	0.714286	1	0.588349	0.411651	0.294037	0.123568633
2	1.071429	0.142857	1	0.497449	0.502551	0.071793	0.005101908
3	0.142857	0.285714	0	0.56092	0.43908	-0.16026	0.020105201
4	0	0	0	0.5	0.5	0	0
5	0.642857	0.714286	1	0.563432	0.436568	0.311834	0.125498161
6	0.071429	0.142857	1	0.533115	0.466885	0.066698	0.005079662
7	0.071429	0.142857	0	0.533115	0.466885	-0.07616	0.005079662
8	0	0	0	0.5	0.5	0	0
9	0.142857	-0.28571	0	0.43908	0.56092	0.125451	0.020105201
10	0.214286	0.142857	1	0.528032	0.471968	0.067424	0.005086004
11	0.071429	0.142857	0	0.533115	0.466885	-0.07616	0.005079662
12	0	0	0	0.5	0.5	0	0
13	0.071429	0.142857	0	0.533115	0.466885	-0.07616	0.005079662
14	0.071429	0.142857	0	0.533115	0.466885	-0.07616	0.005079662
15	0.642857	0.142857	0	0.512752	0.487248	-0.07325	0.005098722
16	0	0	0	0.5	0.5	0	0
17	0	0	0	0.5	0.5	0	0
18	0.357143	0.714286	1	0.612821	0.387179	0.276557	0.121056905
19	0.571429	0.857143	1	0.590818	0.409182	0.350728	0.177613814
20	0	0	0	0.5	0.5	0	0
SUM						1.026371	0.628632858

$$\Delta\theta_s^{\wedge} = 1.6327041687685 \quad \Delta\theta_{s+1}^{\wedge} = -0.63270416876847$$

Dalam proses estimasi parameter, dicari kalkulasi terhadap theta yang paling sesuai untuk peserta. Hasil theta yang paling sesuai dengan peserta akan diperoleh dengan menghitung standart error minimum terhadap pelaksanaan test yang telah dilakukan. Karena kemampuan awal peserta belum diketahui, maka pada mulanya peserta dianggap memiliki tingkat kemampuan sedang. Dalm IRT dengan jangkauan theta berkisar antara 0 dan 1 maka theta dapat diberi nilai 0,5 nilai ini akan

terus berubah ketika tes berlangsung, sedangkan SEM akan semakin menurun hingga berjalannya tes yang semakin banyaknya soal yang terjawab.

Untuk persamaan dari proses perhitungan standart error tingkat kesulitan soal setiap peserta tes dapat dilihat sebagai berikut.

$$SE(\theta_s) = \frac{1}{\sqrt{\sum_{i=1}^N a_i^2 P(\theta^{\wedge}) Q(\theta^{\wedge})}} \dots\dots\dots(3.10)$$

Untuk proses komputasi dapat dilihat sebagai berikut.

$$SE(\theta_s) = \frac{1}{\sqrt{\sum_{i=1}^N a_i^2 P(\theta^{\wedge}) Q(\theta^{\wedge})}}$$

Untuk *item information function* pada model parameter 2 didefinisikan sebagai berikut.

$$I_i(\theta) = a_i^2 p_i(\theta) Q_i(\theta) \dots\dots\dots(3.11)$$

Keterangan:

a_i : merupakan parameter daya beda butir i

$P(\theta)$: $1 / (1 + \text{EXP}(-L))$

$Q_i(\theta)$: $1 - P_i(\theta)$

θ : merupakan *level* tingkat kemampuan minat

Tabel 3. 7 Perhitungan *Item Information Function*

Kemampuan	P(θ)	Q(θ)	P(θ)*Q(θ)	a ²	I(θ)
-3	0.08	0.92	0.0736	0.5102	0.03755
-2	0.14	0.86	0.1204	0.5102	0.06143
-1	0.26	0.74	0.1924	0.5102	0.09816
0	0.41	0.59	0.2419	0.5102	0.12342
1	0.59	0.41	0.2419	0.5102	0.12342

2	0.74	0.26	0.1924	0.5102	0.09816
3	0.86	0.14	0.1204	0.5102	0.06143

Dalam model 2 parameter logistic diasumsikan bahwa dua parameter terhubung dengan pernyataan parameter a_i yang menggambarkan tingkat kesulitan butir soal dan parameter tambahan b_i yang menggambarkan diskriminasi item butir soal. Parameter b menunjukkan sejauhmana butir soal dapat membantu membedakan antara responden yang memiliki kemampuan tinggi dan kemampuan rendah. Kedua parameter tersebut menunjukkan seberapa dekat item terhubung dengan dengan seluruh tes. Semakin tinggi nilai kedua parameter tersebut maka semakin besar diskriminasi item.

BAB IV

UJI COBA DAN PEMBAHASAN

Dalam bab ini akan dibahas mengenai ujicoba dan pembahasan *item response theory* yang diimplementasikan pada aplikasi BEESMART. Diantara yang akan dibahas dalam uji coba pada bab ini adalah, tingkat kesulitan, daya beda butir, nilai probabilitas, *item characteristic curve* dan *item information function*.

4.1 Data Soal Uji coba

Pada bab uji coba dan pembahasan akan dijelaskan mengenai rangkaian uji coba serta evaluasi yang telah dilakukan. Tahap uji coba ini bertujuan untuk mengetahui implementasi *Item response theory* model 2 parameter logistik. Hasil uji coba butir soal terhadap peserta tes diperoleh sebaran jawaban dengan informasi yang disajikan berupa nim, nama, response jawaban sejumlah 29 peserta tes terhadap butir soal 1 sampai dengan 20, keterangan jumlah benar butir soal yang berhasil dijawab, keterangan jumlah butir soal yang tidak berhasil di jawab dan keterangan siswa terhadap test yang telah dilakukan.

Berikut merupakan keterangan untuk masing-masing kolom yang tertera dalam aplikasi *BEESMART*. **No** merupakan fungsi kolom yang menggambarkan jumlah peserta tes yang mengikut ujian. Kolom **Nis** merupakan id ujian yang dimiliki siswa saat melaksanakan ujian, **Nama Siswa** merupakan nama peserta tes yang mengikuti ujian. Kolom **1 sampai dengan 20** merupakan jumlah butir soal pilihan ganda yang harus dijawab oleh peserta tes yang dilambangkan dengan 0 dan 1. Dengan keterangan 0 merupakan keterangan jawaban untuk butir soal yang salah dan 1 merupakan keterangan untuk jawaban butir soal yang benar. Kolom **Benar** digunakan untuk menunjukkan jumlah jawaban benar dari seorang siswa ketika

mengerjakan soal ujian. Kolom **Salah** merupakan jumlah keterangan jawaban salah yang dimiliki siswa saat ujian kolom **Skor total** merupakan fitur bawaan dari aplikasi *BEESMART* yang mampu menampilkan keterangan skor secara keseluruhan. Kolom **Skor keterangan** merupakan fitur bawaan dari aplikasi *BEESMART* yang berfungsi untuk menampilkan keterangan tuntas dan tidak seorang siswa setelah melaksanakan ujian.

Adapun tampilan untuk data soal adalah sebagai berikut dapat dilihat pada Gambar 4.1

No. ▲	Soal ⇅	Pertanyaan ⇅	Jenis ⇅	Level ⇅	Soal ⇅	Opsi ⇅	Lihat ⇅	Edit ⇅	Del ⇅
1	MK1811231046	دَخَلَ بَرَاءٌ مَكْتَبَ فَرَاتٍ لِيَدْعُوهُ...	Pilihan Ganda	Mudah	Tidak	Tidak			
2	MK1811231046	دَخَلَ بَرَاءٌ مَكْتَبَ فَرَاتٍ لِيَدْعُوهُ...	Pilihan Ganda	Mudah	Tidak	Tidak			
3	MK1811231046	دَخَلَ بَرَاءٌ مَكْتَبَ فَرَاتٍ لِيَدْعُوهُ...	Pilihan Ganda	Mudah	Tidak	Tidak			
4	MK1811231046	دَخَلَ بَرَاءٌ مَكْتَبَ فَرَاتٍ لِيَدْعُوهُ...	Pilihan Ganda	Mudah	Tidak	Tidak			
5	MK1811231046	أَخَذَ : هَلْ تَلَّتْ مَنَاسٍ ؟ مَنَعْنِ لَا، أَيْ...	Pilihan Ganda	Mudah	Tidak	Tidak			
6	MK1811231046	أَخَذَ : هَلْ تَلَّتْ مَنَاسٍ ؟ مَنَعْنِ لَا، أَيْ...	Pilihan Ganda	Mudah	Tidak	Tidak			
7	MK1811231046	أَخَذَ : هَلْ تَلَّتْ مَنَاسٍ ؟ مَنَعْنِ لَا، أَيْ...	Pilihan Ganda	Mudah	Tidak	Tidak			
8	MK1811231046	عَمْرُ طَائِبٍ الْيُونَنِيِّ، يَدْرُسُ الْفَتْحَ الْعَرَبِيَّ...	Pilihan Ganda	Mudah	Tidak	Tidak			
9	MK1811231046	عَمْرُ طَائِبٍ الْيُونَنِيِّ، يَدْرُسُ الْفَتْحَ الْعَرَبِيَّ...	Pilihan Ganda	Mudah	Tidak	Tidak			

Gambar 4. 1 Data Soal Ujian

Adapun tampilannya untuk hasil dan analisis soal dapat dilihat pada Gambar 4.2 dan informasi dashboard dapat dilihat pada Gambar 4.3

jumlah kelompok *upper* dan *lower*. Contoh perhitungan hasil komputasi ini diambil dari jawaban benar kelompok *upper* sebanyak 2 butir dan kelompok *lower* 0 butir dibagi dengan jumlah siswa kelompok *upper* dan *lower* sebanyak 18 siswa nilai b diperoleh 0.11111 dan jika mengacu pada pembagian kategori menurut Nitko (1996) maka tingkat kesuliatn butir soal no 3 termasuk pada tingkatan susah. Adapun pembagian kategori tipe butir soal adalah sebagai berikut.

- Jika nilai b diantara 0,00 – 0,30 maka soal tergolong susah
- Jika nilai b diantara 0,31 -- 0,70 maka soal tergolong sedang, dan
- Jika nilai b diantara 0,71 – 1,00 maka soal tergolong mudah.

Selengkapnya perhitungan hasil uji coba tingkat kesulitan dan klasifikasi butir soal dapat dilihat pada Tabel 4.1

Tabel 4. 1 Analisis Butir Untuk Tingkat Kesulitan

Butir Soal	Tingkat Kesulitan	Keterangan
1 (1189)	0.5	SEDANG
2 (1187)	1.1428571428571	MUDAH
3 (1188)	0.21428571428571	SUSAH
4 (1186)	0	SUSAH
5 (1185)	0.64285714285714	SEDANG
6 (1182)	0.14285714285714	SUSAH
7 (1181)	0.14285714285714	SUSAH
8 (1178)	0	SUSAH
9 (1179)	0	SUSAH
10 (1180)	0.14285714285714	SUSAH
11 (1190)	0.071428571428571	SUSAH
12 (1191)	0	SUSAH
13 (1192)	0	SUSAH
14 (1193)	0.071428571428571	SUSAH
15 (1194)	0.5	SEDANG

Butir Soal	Tingkat Kesulitan	Keterangan
16 (1184)	0	SUSAH
17 (1183)	0	SUSAH
18 (1197)	0.5	SEDANG
19 (1196)	0.57142857142857	SEDANG
20 (1195)	0	SUSAH

Tabel 4. 2 Perbandingan data sistem dan wawancara dosen

NO	TINGKAT KESULITAN BUTIR SOAL	
	DOSEN PENGAJAR	SISTEM KE 1
1	SEDANG	SEDANG
2	SEDANG	MUDAH
3	SUSAH	SUSAH
4	SUSAH	SUSAH
5	SEDANG	SEDANG
6	SEDANG	SUSAH
7	SUSAH	SUSAH
8	SUSAH	SUSAH
9	SUSAH	SUSAH
10	SUSAH	SUSAH
11	SEDANG	SUSAH
12	SUSAH	SUSAH
13	SEDANG	SUSAH
14	SUSAH	SUSAH
15	SEDANG	SEDANG
16	MUDAH	SUSAH
17	SUSAH	SUSAH
18	SEDANG	SEDANG
19	SEDANG	SEDANG
20	SUSAH	SUSAH

Pada tabel 4.2 dapat dilihat bahwa perbandingan data yang diperoleh antara data dari sistem dan dosen memiliki kesamaan dari 20 butir soal hanya terdapat 6 butir soal berbeda yaitu pada butir soal no 2, 5, 6, 11, 13, 16.

Tabel 4. 3 Perbandingan Data Sistem 1 dan Papertest

NO	TINGKAT KESULITAN BUTIR SOAL	
	SISTEM KE 1	PAPER TEST KE 2
1	SEDANG	SEDANG
2	SEDANG	MUDAH
3	SUSAH	SUSAH
4	SUSAH	SUSAH
5	SEDANG	SEDANG
6	SUSAH	SUSAH
7	SUSAH	SUSAH
8	SUSAH	SUSAH
9	SUSAH	SUSAH
10	SUSAH	SUSAH
11	SUSAH	SUSAH
12	SUSAH	SUSAH
13	SUSAH	SUSAH
14	SUSAH	SUSAH
15	SEDANG	SEDANG
16	SUSAH	SUSAH
17	SUSAH	SUSAH
18	SEDANG	SUSAH
19	SEDANG	SEDANG
20	SUSAH	SUSAH

Pada Tabel 4.3 dapat dilihat bahwa perbandingan data yang diperoleh antara data dari sistem dan dosen memiliki kesamaan dari 20 butir soal hanya terdapat 2 butir soal berbeda yaitu pada butir soal no 2, 18

4.2.2 Uji Coba Daya Beda Butir Soal

Uji coba perhitungan daya beda merupakan sebuah uji coba yang menyajikan informasi tentang daya beda butir soal. untuk informasi yang disajikan dalam tampilan modal analisis butir soal untuk daya beda ini diantaranya yaitu jumlah butir soal, jumlah kelompok atas, jumlah kelompok bawah, dan daya beda. Selain itu, untuk fungsi dari daya beda adalah untuk mengukur sejauh mana soal

mampu membedakan antara peserta yang belum dan sudah menguasai materi butir soal. Dengan pembagian kemampuan kelompok peserta tes seperti dijelaskan Anastasi dan Urbina (1997) yaitu dalam pembagian kelompok golongan dilakukan dengan mengambil dari 25% urutan nilai terbaik sebagai siswa kelompok atas, dan 25% dari nilai terendah sebagai siswa kelompok bawah. Adapun kelompok peserta tes yang termasuk dalam kelompok atas dan bawah.

Dari *sample* data hasil tes yang diperoleh yaitu jumlah peserta tes sejumlah 29 itu berarti 25% dari 29 peserta merupakan peserta yang memiliki kemampuan tinggi atau rendah. Untuk contoh perhitungan dapat dilihat sebagai berikut. $29 \times 25\% = 7,25$ jika dibulatkan menjadi 7. Sehingga dari 29 peserta tes 7 nilai tertinggi tergolong dalam peserta yang memiliki kemampuan tinggi dan 7 peserta tes yang memiliki nilai terendah adalah peserta yang tergolong dalam peserta yang memiliki kemampuan rendah. Untuk contoh perhitungan diambil dari butir soal no 3 yaitu peserta yang menjawab benar sebanyak pada kelompok atas sebanyak 6, dan 1

jawaban benar dari kelompok peserta yang tergolong dalam kemampuan rendah. Selengkapnya hasil uji coba daya beda butir soal dapat dilihat pada Tabel 4.5

Tabel 4. 4 Analisis Butir Untuk Daya Beda

Butir Soal	Kelompok Atas	Kelompok Bawah	Daya Beda(a)	Keterangan
1	6	1	0.714285714	Butir Soal Baik/Terima
2	8	7	0.142857143	Butir Soal Ditolak

Butir Soal	Kelompok Atas	Kelompok Bawah	Daya Beda(a)	Keterangan
3	2	0	0.285714286	Butir Soal Perlu Diperbaiki
4	0	0	0	Butir Soal Ditolak
5	7	2	0.714285714	Butir Soal Baik/Terima
6	1	0	0.142857143	Butir Soal Ditolak
7	1	0	0.142857143	Butir Soal Ditolak
8	0	0	0	Butir Soal Ditolak
9	0	2	- 0.285714286	Butir Soal Ditolak
10	2	1	0.142857143	Butir Soal Ditolak
11	1	0	0.142857143	Butir Soal Ditolak
12	0	0	0	Butir Soal Ditolak
13	1	0	0.142857143	Butir Soal Ditolak
14	1	0	0.142857143	Butir Soal Ditolak
15	5	4	0.142857143	Butir Soal Ditolak
16	0	0	0	Butir Soal Ditolak
17	0	0	0	Butir Soal Ditolak
18	5	0	0.714285714	Butir Soal Baik/Terima
19	7	1	0.857142857	Butir Soal Baik/Terima
20	0	0	0	Butir Soal Ditolak

Dari perolehan nilai daya beda pada tabel 4.5 maka nilai daya beda menurut Crocker dan Algina (1986), dikalkulasikan sebagai berikut:

- a. Jika daya pembeda $> 0,4$ maka butir soal termasuk dalam kategori baik
- b. Jika nilai daya pembeda berkisar antara $0,3 >$ sampai dengan $< 0,4$ maka butir soal termasuk dalam kategori cukup baik.
- c. Jika nilai daya pembeda berkisar antara $0,2 >$ sampai dengan $\neg < 0,3$ maka butir soal termasuk dalam kategori kurang baik atau perlu diperbaiki.
- d. Jika daya pembeda $< 0,2$ maka butir soal ditolak.

Parameter tingkat kesulitan, faktor daya pembeda, dan faktor tebakan (*guessing*) ini dapat digunakan sebagai parameter pembantu dalam melakukan pendekatan yang dilakukan dengan *Item response theory*. Untuk lebih jelasnya dapat dilihat pada LAMPIRAN.

4.3 Uji Coba Analisis Butir Soal Menggunakan IRT Model 2 Parameter Logistic.

Uji coba analisis butir soal menggunakan IRT model 2 parameter logistic ini dilakukan dengan menentukan *Item characteristic curve* (icc), *standart error* (se) dan *item information function* (IIF).

4.3.1 Uji Coba *Item Characteristic Curve* (ICC).

Model logistik 2 parameter merupakan model parameter yang melibatkan parameter tingkat kesulitan *item* dan parameter daya beda butir soal. Tahap pertama dalam melakukan perhitungan *logistic function* adalah dengan cara menentukan penyimpangan *logistic*, adapun persamaannya dapat dilihat pada persamaan 3.4

sampai dengan 3.7. Tahap tersebut merupakan sebuah langkah untuk memperoleh nilai probabilitas untuk 1 butir soal. Tahap selanjutnya yaitu menentukan fungsi logistic pada setiap response sebaran butir jawaban siswa dengan menggunakan parameter tingkat kesulitan yang dilambangkan dengan b untuk didefinisikan sebagai nilai pada kemampuan probabilitas response yang memiliki nilai khas yang berkisar $-3 \leq b \leq +3$ sehingga dari pengukuran dengan titik khas tersebut item characteristic curve akan terbentuk sebuah curve s yang merupakan gambaran kemampuan siswa.

Untuk hasil uji coba model 2 parameter logistic dengan tingkatan level -3 sampai dengan level 3 untuk butir soal no 3 dapat dilihat pada Tabel 4.6

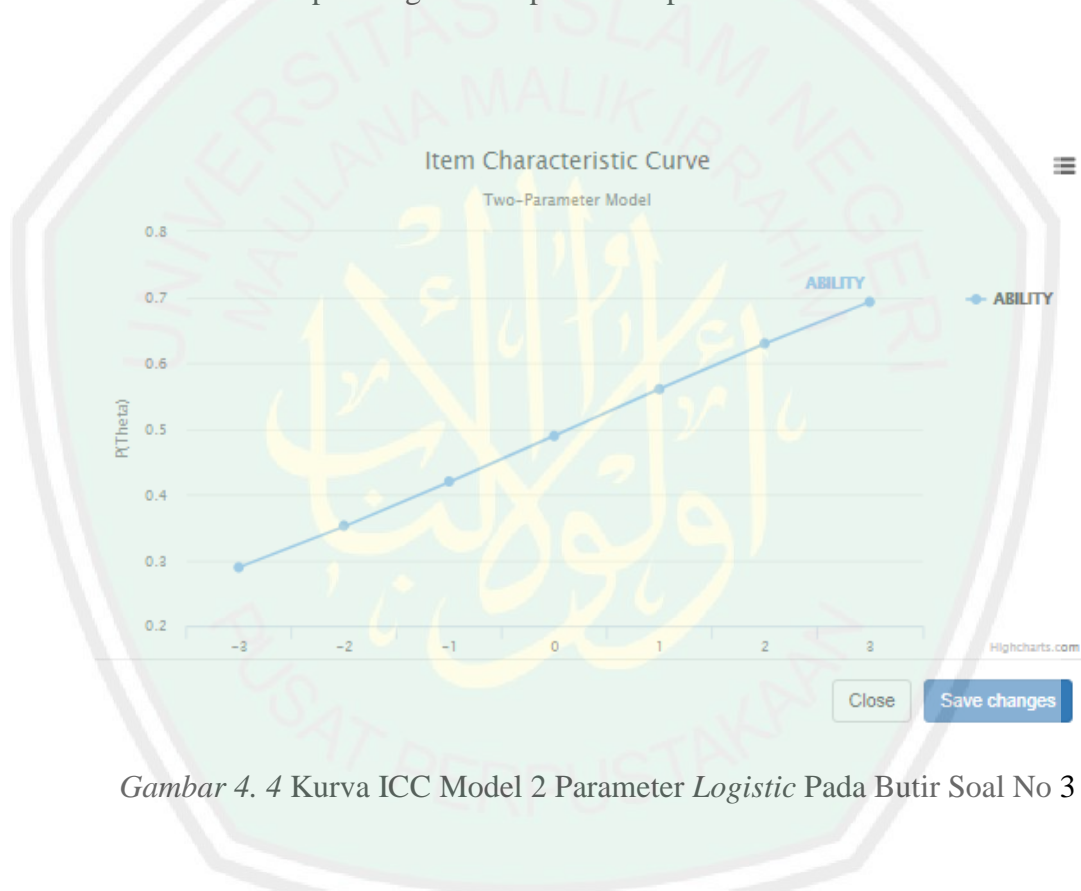
Tabel 4. 5 Perhitungan Model 2 Parameter Logistic Dengan Tingkatan Level -3 Sampai Dengan Level 3 Untuk Butir Soal No 3

Kemampuan	$L=a*(\theta - b)$	$EXP(-L)$	$1 + EXP(-L)$	P
-3	-0.897959184	2.45459	3.454588632	0.289470066
-2	-0.612244898	1.84457	2.84456762	0.351547277
-1	-0.326530612	1.38615	2.386150682	0.419085017
0	-0.040816327	1.04166	2.041660763	0.489797335
1	0.244897959	0.78278	1.78278441	0.560920319
2	0.530612245	0.58824	1.58824471	0.629625897
3	0.816326531	0.44205	1.442052542	0.693456009

Dari hasil uji coba butir soal no 3 diperoleh nilai probabilitas yang berbeda disetiap level kemampuan siswa. Perbedaan nilai probabilitas pada setiap level tersebut akan membentuk sebuah titik yang berbeda disetiap level kemampuannya. Sehingga jika ditarik garis lurus terhadap titik satu dan lainnya maka akan membentuk sebuah kurva ICC. Pada dasarnya ICC adalah penggambaran dalam bentuk kurva yang menjelaskan hubungan antara kemampuan dengan *response*

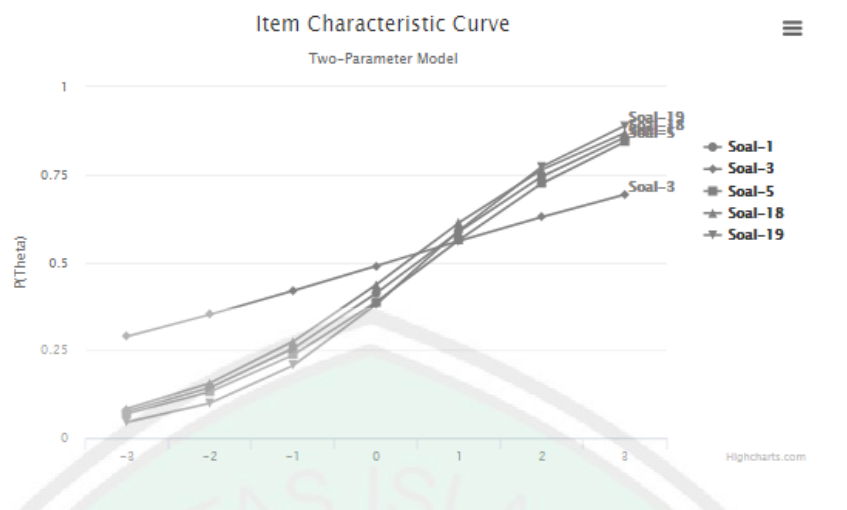
sebuah item. Semakin curam kemiringan nilai dari kurva maka semakin besar nilai daya beda butir soal terhadap siswa, atau dengan kata lain semakin curam kurva maka semakin mudah untuk mengetahui karakter daya beda butir soal terhadap kemampuan siswa.

Adapun tampilan kurva uji coba untuk *Item Characteristic Curve* (ICC) butir 3 dapat dilihat pada Gambar 4.4 dan untuk tampilan kurva perhitungan ICC untuk butir soal no 1 sampai dengan 20 dapat dilihat pada Gambar 4.5



Gambar 4. 4 Kurva ICC Model 2 Parameter Logistic Pada Butir Soal No 3

Gambar 4.4 menunjukkan hasil kurva dari kemampuan siswa terhadap butir soal no 3 dengan tingkat kesulitan butir soal no 3 berada pada -3 sampai dengan +3. Untuk mengetahui dan membandingkan informasi tingkat kesulitan yang dimiliki setiap butir soal terhadap kemampuan siswa dapat dilihat pada Gambar 4.5.



Gambar 4. 5 Kurva ICC Model 2 Parameter Logistic Pada Butir Soal No 1 Sd 20

Gambar 4.5 menunjukkan hasil kurva dari kemampuan siswa terhadap butir soal dari no 1 sampai dengan no 20, dengan tingkat kesulitan butir soal dari -3 sampai dengan +3 . Dalam Gambar 4.5 posisi kurva menunjukkan tingkat kesulitan butir soal, semakin kekiri posisi kurva yang dimiliki butir soal maka semakin susah tingkatan butir soal terhadap kemampuan siswa. Begitu pula sebaliknya semakin kekanan posisi kurva butir soal maka semakin mudah tingkat kesulitan butir soal terhadap kemampuan siswa. Sehingga dapat disimpulkan bahwasanya butir soal tersulit pada informasi kurva Gambar 4.5 dimiliki oleh butir soal no 3, dan butir soal termudah dimiliki oleh butir soal no 19. Pada Gambar 4.5 Kurva yang ditampilkan ada sejumlah 4 yang menggambarkan jumlah butir soal yang diterima dengan kategori butir soal baik dan butir soal cukup baik. Analisis model logistic dua parameter menghasilkan daya beda yang berbeda untuk setiap butir soalnya. Hal ini dapat dilihat pada Tabel 4.6 dimana daya beda butir soal terkecil bernilai 0 dan daya beda butir soal terbesar bernilai parameter daya beda bergerak antara 0 sd 2 (Hambleton, dkk,1991). Daya beda butir soal yang terbentuk pada model ICC

dapat dilihat dari kemiringan kurva yaitu jika butir soal digambarkan dengan kurva yang landai ini berarti nilai daya beda butir soal adalah kecil tetapi jika butir soal digambarkan dengan kurva yang curam berarti butir soal memiliki daya beda yang besar. Jika dilihat pada Gambar 4.5 dari kurva yang ditampilkan maka daya beda terbesar dapat dilihat pada butir soal no 19 dengan nilai daya beda adalah 0.8571 dan untuk butir no 1, 5, 18 dan 19 memiliki kecuraman yang sama, hal itu dapat dilihat dari nilai daya beda yang ke tiga butir soal tersebut memiliki nilai daya beda sebesar 0.7142 Adapun untuk rincian keterangan daya beda dapat dilihat pada Tabel 4.6.

Tabel 4. 6 Daya Beda Butir Soal No 1 s/d 20

Butir Soal	Tingkat Kesulitan(b)	Daya Beda
1	0.5	0.714285714
2	1.071428571	0.142857143
3	0.142857143	0.285714286
4	0	0
5	0.642857143	0.714285714
6	0.071428571	0.142857143
7	0.071428571	0.142857143
8	0	0
9	0.142857143	-0.285714286
10	0.214285714	0.142857143
11	0.071428571	0.142857143
12	0	0
13	0.071428571	0.142857143
14	0.071428571	0.142857143
15	0.642857143	0.142857143
16	0	0
17	0	0
18	0.357142857	0.714285714
19	0.571428571	0.857142857
20	0	0

4.3.2 Uji Coba *Standart Error* (Se).

Pada uji coba iterasi tingkat kesulitan *response* jawaban siswa yang sudah diberikan terhadap butir soal akan dilakukan estimasi tingkat kesulitan butir soal guna memperkirakan tingkat kemampuan masing-masing siswa terhadap butir soal. Untuk persamaan dari proses perhitungan estimasi tingkat kesulitan butir soal dapat dilihat pada persamaan 3.9. Untuk contoh perhitungan dapat diambil dari perkiraan kemampuan nanik sofiatun yang merupakan salah satu peserta tes yang mengikuti tes kemampuan siswa. Adapun persamaan dari proses perhitungan standart error tingkat kesulitan soal setiap peserta tes dapat dilihat pada persamaan 3.10. Untuk tampilan halaman modal analisis butir untuk literasi dapat dilihat pada Gambar 4.6 dan untuk hasil literasi kemampuan siswa terhadap butir soal dapat di ketahui pada Gambar 4.7

7	022	Ali Murtadho	SE
8	028	Sondwi Haris	SE
9	018	M Farul Bari	SE
10	029	Heri Lestari	SE
11	010	Muhammad Alfandoli	SE
12	026	Izul Azmi	SE
13	011	Fajar Santosa	SE
14	027	Illa Hidayati	SE
15	012	Muhammad Sirojui Umam	SE
16	017	Widan Ansori	SE
17	013	M Fajar	SE
18	019	Hakim Muhammad	SE
19	016	Ali Umar Ritonga	SE
20	021	Wahyu Alif	SE
21	023	Dimas Faisal	SE
22	014	Abdul Fatah	SE

Gambar 4. 6 Analisis Butir Untuk Iterasi

Hasil Iterasi ke-20

Item	b	a	U(response)	p	q	a*(u-p)	a*a*(p*q)
1	0.5	0.71428571428571	1	0.8	0.2	0.14285714285714	0.081632653061223
2	1.0714285714286	0.14285714285714	1	0.55	0.45	0.064285714285713	0.0050510204081631
3	0.14285714285714	0.28571428571429	1	0.66	0.34	0.097142857142859	0.018318367346939
4	0	0	0	0.5	0.5	-0	0
5	0.64285714285714	0.71428571428571	0	0.79	0.21	-0.56428571428571	0.084642857142856
6	0.071428571428571	0.14285714285714	0	0.59	0.41	-0.084285714285713	0.0049367346938774
7	0.071428571428571	0.14285714285714	0	0.59	0.41	-0.084285714285713	0.0049367346938774
8	0	0	0	0.5	0.5	-0	0
9	0.14285714285714	-0.28571428571429	0	0.34	0.66	0.097142857142859	0.018318367346939
10	0.21428571428571	0.14285714285714	0	0.58	0.42	-0.082857142857141	0.0049714285714284
11	0.071428571428571	0.14285714285714	1	0.59	0.41	0.058571428571427	0.0049367346938774
12	0	0	0	0.5	0.5	-0	0
13	0.071428571428571	0.14285714285714	0	0.59	0.41	-0.084285714285713	0.0049367346938774
14	0.071428571428571	0.14285714285714	0	0.59	0.41	-0.084285714285713	0.0049367346938774
15	0.64285714285714	0.14285714285714	1	0.57	0.43	0.06142857142857	0.0050020408163263
16	0	0	0	0.5	0.5	-0	0
17	0	0	0	0.5	0.5	-0	0
18	0.35714285714286	0.71428571428571	1	0.82	0.18	0.12857142857143	0.075306122448979
19	0.57142857142857	0.85714285714286	0	0.84	0.16	-0.72	0.098742857142858
20	0	0	0	0.5	0.5	-0	0
SUM						-1.0542857142857	0.4166693877551

Delta Theta Soal Ke-20 = -2.5302691900788

Theta (S+1) Soal Ke-20 = 3.5302691900788

SE dari Arya Bayu = 1.5491882799173

Gambar 4. 7 Hasil Literasi Kemampuan Siswa Terhadap Butir Soal.

4.3.3 Uji Coba Item Information Function (Iif).

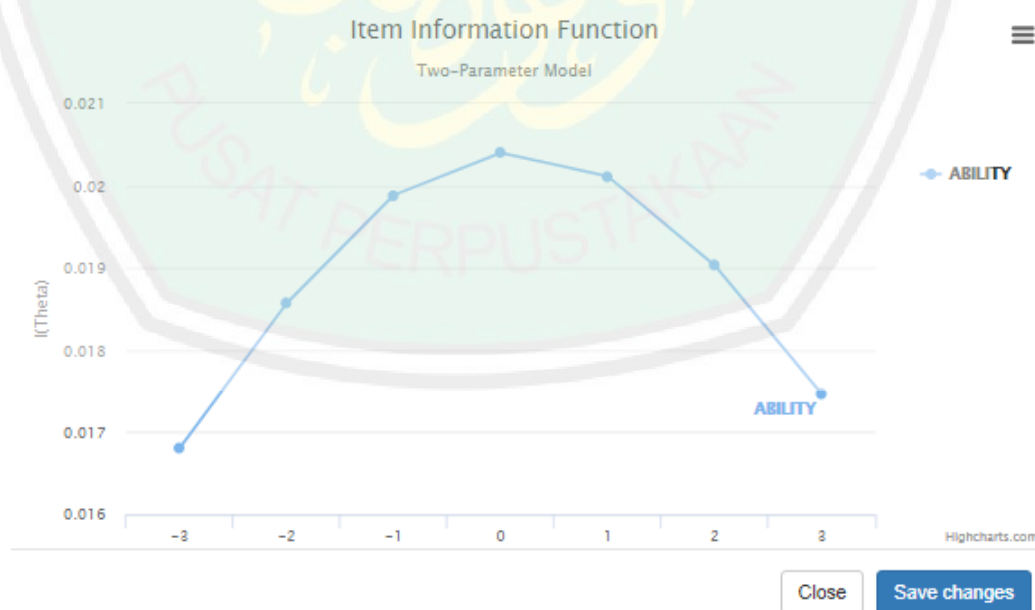
Pada uji coba IIF didefinisikan oleh *fisher* sebagai definisi dari hubungan timbal balik dari keakuratan item soal yang telah diestimasi. Dilakukan juga proses perhitungan IIF atau *Item Information Function* yang bertujuan untuk memperoleh informasi keseluruhan dari tes yang telah dilaksanakan. Untuk *item information function* pada model parameter 2 didefinisikan pada persamaan 3.11.

Hasil uji coba perhitungan IIF dapat dilihat pada Tabel 4.7 menunjukkan hasil uji coba perhitungan kemampuan siswa terhadap butir soal no 3 dengan tingkat kesulitan butir soal no 3 berada pada -3 sampai dengan +3. Dengan *Item informasi function*.

Tabel 4. 7 Perhitungan IIF

Kemampuan	P(θ)	Q(θ)	P(θ)*Q(θ)	a ²	I(θ)
-3	0.29	0.71	0.2059	0.081632653	0.016808163
-2	0.35	0.65	0.2275	0.081632653	0.018571429
-1	0.42	0.58	0.2436	0.081632653	0.019885714
0	0.49	0.51	0.2499	0.081632653	0.0204
1	0.56	0.44	0.2464	0.081632653	0.020114286
2	0.63	0.37	0.2331	0.081632653	0.019028571
3	0.69	0.31	0.2139	0.081632653	0.017461224

Adapun untuk isi informasi IIF bisa dilihat pada Gambar 4.8 Fungsi informasi butir soal untuk butir soal no 3 terhadap kemampuan siswa dapat diartikan bahwa plot tersebut secara umum dapat memberikan informasi bahwa tes dapat mengukur dengan baik, Hal ini dikarenakan pada Tabel 4 terlihat bahwa dayabeda pada semua item berkisar antara 0 sampai dengan 1. Maka dengan kemampuan sedikit dibawah nol dengan fungsi informasi 0.00475 ini berarti tes dapat mendiskriminasi dengan baik pada kemampuan tersebut.



Gambar 4. 8 Gambar Item Information Function

4.3.4 Uji Coba Tingkat Kesulitan, Daya beda, ICC dan IIF.

Uji coba tingkat kesulitan, daya beda, ICC dan IIF merupakan tabel perbandingan nilai daya beda butir soal dan tingkat kesulitan. adapun untuk ICC dan IIF merupakan grafik nilai yang diperoleh dari perhitungan lanjut dari tingkat kesulitan dan daya beda yang dapat dijelaskan dalam bentuk grafik.

Table 4.8 Uji Coba Tingkat Kesulitan dan Dayabeda

Butir Soal	Tingkat kesulitan	Keterangan Butir	Daya Beda	Keterangan	ICC
1	0.5	SEDANG	0.714285714	Butir Soal Baik/Terima	Pada butir soal no 1 diketahui bahwa nilai tingkat kesulitan adalah 0.5 dan nilai dayabeda adalah 0.714285714 sehingga dari nilai tingkat kesulitan dan daya beda yang akan dihitung dengan nilai khas antara -3 s/d 3 maka akan terbentuk sebuah curva yang dapat dilihat pada Lampiran
2	1.0714285714286	MUDAH	0.142857143	Butir Soal Ditolak	Pada butir soal no 2 diketahui bahwa nilai tingkat kesulitan adalah 1.0714285714286 dan nilai dayabeda adalah 0.142857143 sehingga dari nilai tingkat kesulitan dan daya beda yang akan dihitung dengan nilai khas antara -3 s/d 3 maka akan terbentuk sebuah curva yang dapat dilihat pada Lampiran

Butir Soal	Tingkat kesulitan	Keterangan Butir	Daya Beda	Keterangan	ICC
3	0.14285714285714	SUSAH	0.285714286	Butir Soal Perlu Diperbaiki	Pada butir soal no 3 diketahui bahwa nilai tingkat kesulitan adalah 0.14285714285714 dan nilai dayabeda adalah 0.285714286 sehingga dari nilai tingkat kesulitan dan daya beda yang akan dihitung dengan nilai khas antara -3 s/d 3 maka akan terbentuk sebuah curva yang dapat dilihat pada Lampiran
4	0	SUSAH	0	Butir Soal Ditolak	Pada butir soal no 4 diketahui bahwa nilai tingkat kesulitan adalah 0 dan nilai dayabeda adalah 0 sehingga dari nilai tingkat kesulitan dan daya beda yang akan dihitung dengan nilai khas antara -3 s/d 3 maka akan terbentuk sebuah curva yang dapat dilihat pada Lampiran
5	0.64285714285714	SEDANG	0.714285714	Butir Soal Baik/Terima	Pada butir soal no 5 diketahui bahwa nilai tingkat kesulitan adalah 0.64285714285 dan nilai dayabeda adalah 0.714285714 sehingga dari nilai tingkat kesulitan dan daya beda yang akan dihitung dengan nilai khas antara -3 s/d 3 maka akan terbentuk sebuah curva yang dapat dilihat pada Lampiran.

Butir Soal	Tingkat kesulitan	Keterangan Butir	Daya Beda	Keterangan an	ICC
6	0.0714 285714 28571	SUSAH	0.1428 57143	Butir Soal Ditolak	Pada butir soal no 6 diketahui bahwa nilai tingkat kesulitan adalah 0.071428571428571 dan nilai dayabeda adalah 0.142857143 sehingga dari nilai tingkat kesulitan dan daya beda yang akan dihitung dengan nilai khas antara -3 s/d 3 maka akan terbentuk sebuah curva yang dapat dilihat pada Lampiran.
7	0.0714 285714 28571	SUSAH	0.1428 57143	Butir Soal Ditolak	Pada butir soal no 7 diketahui bahwa nilai tingkat kesulitan adalah 0.07142857142 dan nilai dayabeda adalah 0.142857143 sehingga dari nilai tingkat kesulitan dan daya beda yang akan dihitung dengan nilai khas antara -3 s/d 3 maka akan terbentuk sebuah curva yang dapat dilihat pada Lampiran.
8	0	SUSAH	0	Butir Soal Ditolak	Pada butir soal no 8 diketahui bahwa nilai tingkat kesulitan adalah 0 dan nilai dayabeda adalah 0 sehingga dari nilai tingkat kesulitan dan daya beda yang akan dihitung dengan nilai khas antara -3 s/d 3 maka akan terbentuk sebuah curva yang dapat dilihat pada Lampiran.

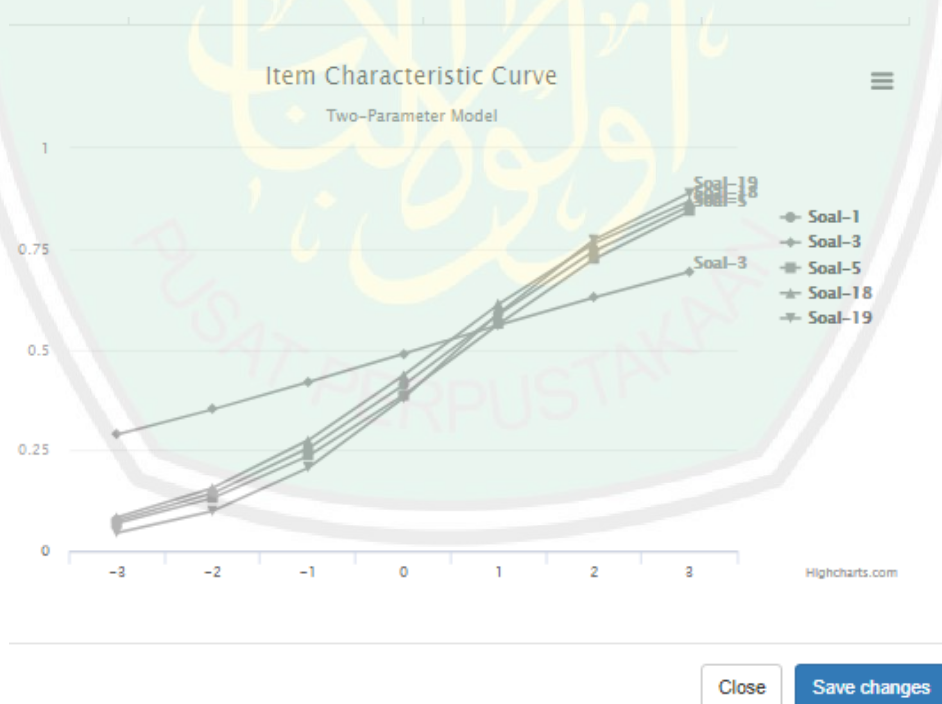
Butir Soal	Tingkat kesulitan	Keterangan Butir	Daya Beda	Keterangan an	ICC
9	0.1428 571428 5714	SUSAH	0.2857 14286	Butir Soal Ditolak	Pada butir soal no 9 diketahui bahwa nilai tingkat kesulitan adalah 0.14285714285714 dan nilai dayabeda adalah 0.285714286 sehingga dari nilai tingkat kesulitan dan daya beda yang akan dihitung dengan nilai khas antara -3 s/d 3 maka akan terbentuk sebuah curva yang dapat dilihat pada Lampiran.
10	0.2142 857142 8571	SUSAH	0.1428 57143	Butir Soal Ditolak	Pada butir soal no 10 diketahui bahwa nilai tingkat kesulitan adalah 0.214285714 dan nilai dayabeda adalah 0.142857143 sehingga dari nilai tingkat kesulitan dan daya beda yang akan dihitung dengan nilai khas antara -3 s/d 3 maka akan terbentuk sebuah curva yang dapat dilihat pada Lampiran.
11	0.0714 285714 28571	SUSAH	0.1428 57143	Butir Soal Ditolak	Pada butir soal no 11 diketahui bahwa nilai tingkat kesulitan adalah 0.071428571 dan nilai dayabeda adalah 0.142857143 sehingga dari nilai tingkat kesulitan dan daya beda yang akan dihitung dengan nilai khas antara -3 s/d 3 maka akan terbentuk sebuah curva yang dapat dilihat pada Lampiran.

Butir Soal	Tingkat kesulitan	Keterangan Butir	Daya Beda	Keterangan	ICC
12	0	SUSAH	0	Butir Soal Ditolak	Pada butir soal no 12 diketahui bahwa nilai tingkat kesulitan adalah 0 dan nilai dayabeda adalah 0 sehingga dari nilai tingkat kesulitan dan daya beda yang akan dihitung dengan nilai khas antara -3 s/d 3 maka akan terbentuk sebuah curva yang dapat dilihat pada Lampiran.
13	0.0714 285714 28571	SUSAH	0.1428 57143	Butir Soal Ditolak	Pada butir soal no 13 diketahui bahwa nilai tingkat kesulitan adalah 0.071428571428571 dan nilai dayabeda adalah 0.142857143 sehingga dari nilai tingkat kesulitan dan daya beda yang akan dihitung dengan nilai khas antara -3 s/d 3 maka akan terbentuk sebuah curva yang dapat dilihat pada Lampiran.
14	0.0714 285714 28571	SUSAH	0.1428 57143	Butir Soal Ditolak	Pada butir soal no 14 diketahui bahwa nilai tingkat kesulitan adalah 0.071428571428571 dan nilai dayabeda adalah 0.142857143 sehingga dari nilai tingkat kesulitan dan daya beda yang akan dihitung dengan nilai khas antara -3 s/d 3 maka akan terbentuk sebuah curva yang dapat dilihat pada Lampiran.

Butir Soal	Tingkat kesulitan	Keterangan Butir	Daya Beda	Keterangan	ICC
15	0.64285714285714	SEDANG	0.142857143	Butir Soal Ditolak	Pada butir soal no 15 diketahui bahwa nilai tingkat kesulitan adalah 0.64285714285714 dan nilai dayabeda adalah 0.142857143 sehingga dari nilai tingkat kesulitan dan daya beda yang akan dihitung dengan nilai khas antara -3 s/d 3 maka akan terbentuk sebuah curva yang dapat dilihat pada Lampiran.
16	0	SUSAH	0	Butir Soal Ditolak	Pada butir soal no 16 diketahui bahwa nilai tingkat kesulitan adalah 0 dan nilai dayabeda adalah 0 sehingga dari nilai tingkat kesulitan dan daya beda yang akan dihitung dengan nilai khas antara -3 s/d 3 maka akan terbentuk sebuah curva yang dapat dilihat pada Lampiran.
17	0	SUSAH	0	Butir Soal Ditolak	Pada butir soal no 17 diketahui bahwa nilai tingkat kesulitan adalah 0 dan nilai dayabeda adalah 0 sehingga dari nilai tingkat kesulitan dan daya beda yang akan dihitung dengan nilai khas antara -3 s/d 3 maka akan terbentuk sebuah curva yang dapat dilihat pada Lampiran.

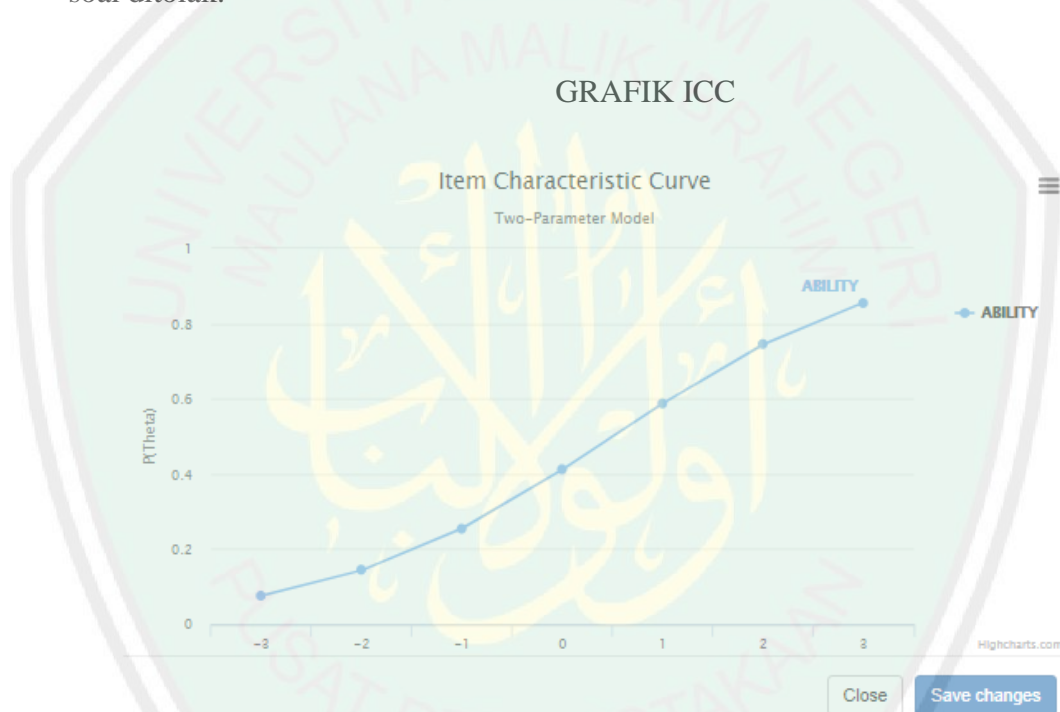
Butir Soal	Tingkat kesulitan	Keterangan Butir	Daya Beda	Keterangan	ICC
18	0.3571 428571 4286	SEDANG	0.7142 85714	Butir Soal Baik/Terima	Pada butir soal no 18 diketahui bahwa nilai tingkat kesulitan adalah 0.3571428571 dan nilai dayabeda adalah 0.714285714 sehingga dari nilai tingkat kesulitan dan daya beda yang akan dihitung dengan nilai khas antara -3 s/d 3 maka akan terbentuk sebuah curva yang dapat dilihat pada Lampiran.
19	0.5714 285714 2857	SEDANG	0.8571 42857	Butir Soal Baik/Terima	Pada butir soal no 19 diketahui bahwa nilai tingkat kesulitan adalah 0.571428571 dan nilai dayabeda adalah 0.857142857 sehingga dari nilai tingkat kesulitan dan daya beda yang akan dihitung dengan nilai khas antara -3 s/d 3 maka akan terbentuk sebuah curva yang dapat dilihat pada Lampiran.
20	0	SUSAH	0	Butir Soal Ditolak	Pada butir soal no 20 diketahui bahwa nilai tingkat kesulitan adalah 0 dan nilai dayabeda adalah 0 sehingga dari nilai tingkat kesulitan dan daya beda yang akan dihitung dengan nilai khas antara -3 s/d 3 maka akan terbentuk sebuah curva yang dapat dilihat pada Lampiran.

Adapun tabel diatas merupakan keterangan singkat yang diperoleh dari butir soal no 1 sampai dengan butir soal no 20. Yaitu hasil dari perhitungan ICC adalah dalam bentuk grafik. Kemiringan kurva yang terbentuk pada ICC pemunjukan tingkat kesulitan yang dimiliki butir soal terhadap kemampuan siswa dan kecuraman kurva yang terbentuk pada ICC menunjukkan besaran daya beda yang dimiliki butir soal. Namun tingkat kemiringan dan kecuraman sebuah butir soal masih belum bisa di gambarkan secara pasti jika semua butir soal belum memiliki nilai tingkat kesuslitan dan daya beda. Oleh karenanya grafik ICC dan IIF baru akan nampak secara keseluruhan ketika test telah selesai berlangsung, dan semua butir soal memiliki nilai daya beda dan tingkat kesulitan yang diperoleh dari jawaban peserta tes yang telah mengikuti tes.



Gambar 4. 9 *Item Characteristic Curve*

Analisis model logistic dua parameter menghasilkan daya beda yang berbeda untuk setiap butir soalnya. Hal ini dapat dilihat pada Tabel 4.8 dimana daya beda butir soal terkecil bernilai 0 dan daya beda butir soal terbesar bernilai 0.41379. Parameter daya beda bergerak antara 0 sd 2 (Hambleton, dkk,1991). Dengan klasifikasi daya beda menurut Crocker dan Algina (1986) maka dari 20 butir soal yang diujikan kepada peserta tes hanya 4 butir tes yang dapat digunakan dengan rincian 3 butir soal bernilai cukup baik ,1 soal bernilai baik dan sisanya butir soal ditolak.



Gambar 4. 10 *Item Characteristic Curve* Butir Soal No 1

Pada butir soal no 1 merupakan tipe butir soal dengan kategori tingkat kesulitan yang sedang, hal itu dibuktikan dengan nilai tingkat kesulitan pada butir no 1 adalah 0,5. Selain nilai tingkat kesulitan, pada butir soal no 1 juga memiliki nilai daya beda sebesar 0.714285714 yang pada kasus ini berarti butir soal no 1 berkategori soal yang cukup baik. Grafik ICC butir soal no 1 terbentuk dari nilai

probabilitas yang bernilai khas antara -3 sampai dengan 3 untuk lebih jelasnya nilai ICC butir soal no 1 dapat dilihat pada Tabel 4.9.

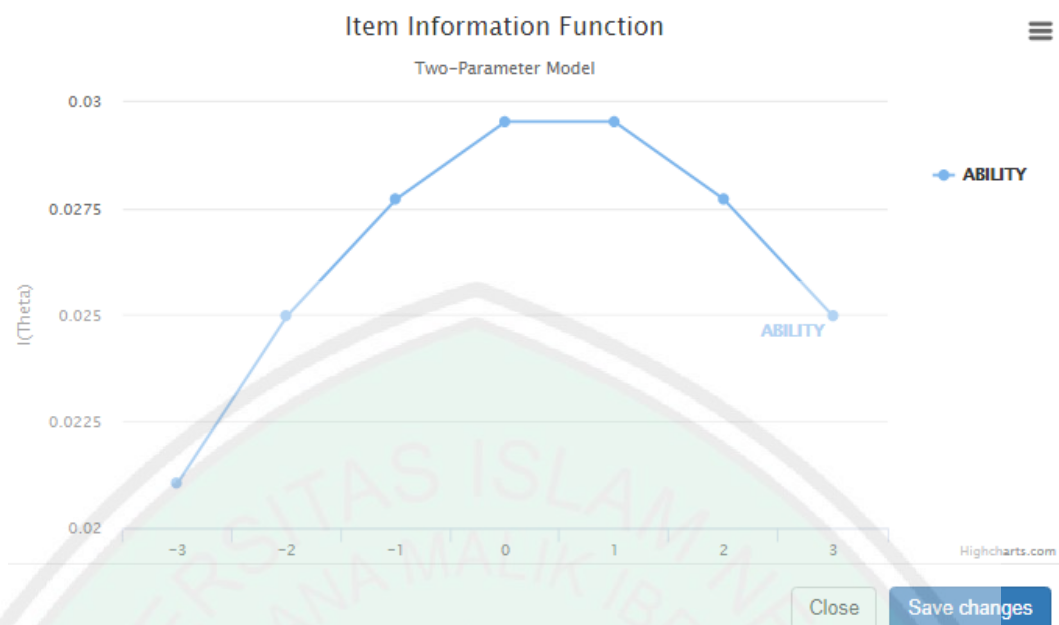
Tabel 4.9 Nilai ICC Butir 1

Kemampuan	$L=a*(\theta - b)$	$EXP(-L)$	$1 + EXP(-L)$	P
-3	-2.5	12.182494	13.182494	0.07585818
-2	-1.78571429	5.96383831	6.96383831	0.14359897
-1	-1.07142857	2.9195473	3.9195473	0.2551315
0	-0.35714286	1.42924003	2.42924003	0.41165138
1	0.35714286	0.69967254	1.69967254	0.58834862
2	1.07142857	0.34251886	1.34251886	0.7448685
3	1.78571429	0.16767725	1.16767725	0.85640103

Nilai *Item Information Function* (IIF) pada butir soal no 1 diperoleh dari nilai Informasi butir tes yang bernilai khas antara -3 sampai dengan 3, untuk lebih jelasnya nilai informasi butir soal no1 dapat dilihat pada table 4.10. dari hasil informasi butir tes dengan nilai khas -3 sampai dengan 3 pada butir soal no 3 terbentuklah sebuah kurva IIF butir soal no 1 yang dapat dilihat pada Gambar 5.2.

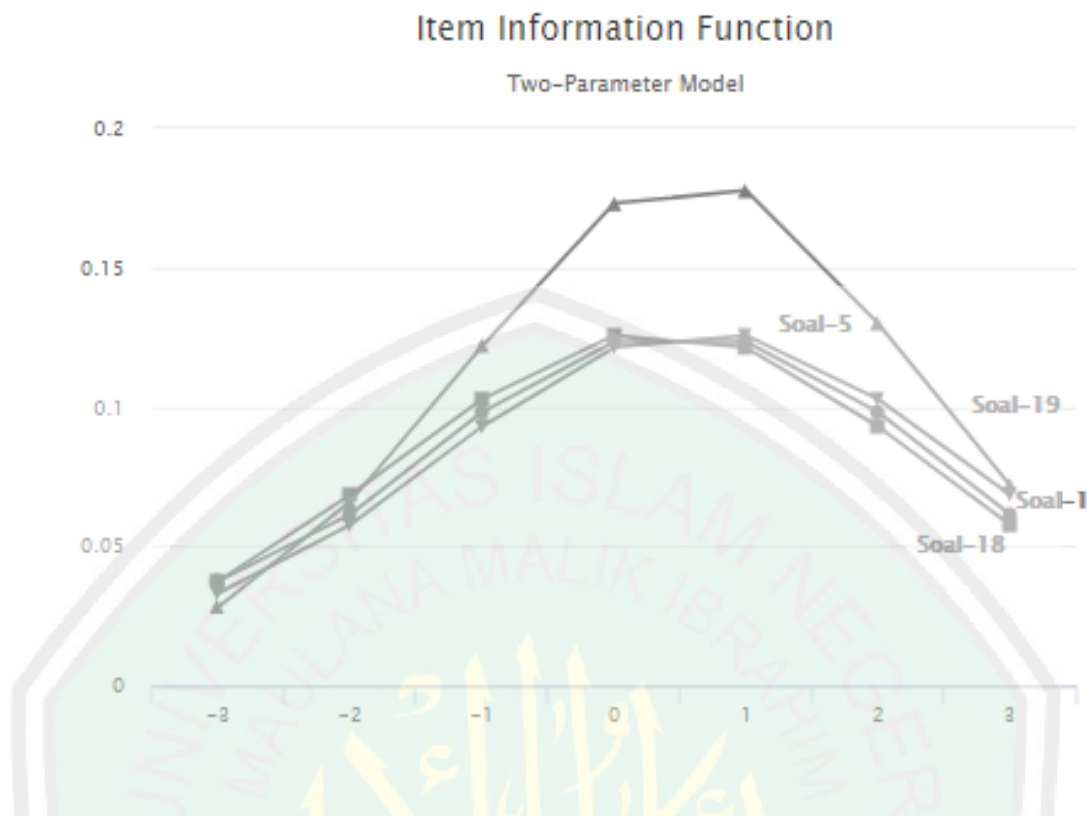
Tabel 4.10 Nilai IIF Butir Soal No 1

Kemampuan	$P(\theta)$	$Q(\theta)$	$P(\theta)*Q(\theta)$	a^2	$I(\theta)$
-3	0.08	0.92	0.0736	0.51020408	0.03755
-2	0.14	0.86	0.1204	0.51020408	0.06143
-1	0.26	0.74	0.1924	0.51020408	0.09816
0	0.41	0.59	0.2419	0.51020408	0.12342
1	0.59	0.41	0.2419	0.51020408	0.12342
2	0.74	0.26	0.1924	0.51020408	0.09816
3	0.86	0.14	0.1204	0.51020408	0.06143



Gambar 4. 11 *Item Information Function* Butir Soal No 1

Adapun hubungan tingkat kesulitan dan daya beda pada ICC dan IIF adalah dalam bentuk grafik. Kemiringan kurva yang terbentuk pada ICC pemunculan tingkat kesulitan yang dimiliki butir soal terhadap kemampuan siswa dan kecuraman kurva yang terbentuk pada ICC menunjukkan besaran daya beda yang dimiliki butir soal. Namun tingkat kemiringan dan kecuraman sebuah butir soal masih belum bisa di gambarkan secara pasti jika semua butir soal belum memiliki nilai tingkat kesulitan dan daya beda. Oleh karenanya grafik ICC dan IIF baru akan nampak secara keseluruhan ketika test telah selesai berlangsung, dan semua butir soal memiliki nilai daya beda dan tingkat kesulitan yang diperoleh dari jawaban peserta tes yang telah mengikuti tes.



Gambar 4. 12 Tampilan Item Information Function Butir 1 sampai dengan 20

4.4 Integrasi Evaluasi Pembelajaran Dalam Islam

Belajar adalah sebuah keutamaan bagi setiap orang manusia, hal ini dikarenakan dengan belajar seorang akan menjadi tahu selain itu memiliki pengetahuan yang luas akan sebuah ilmu akan di tingkatkan derajatnya oleh Alloh SWT, sebagai mana yang telah di firmankan Alloh SWT pada Q.S Al-mujadalah ayat 11;

يَا أَيُّهَا الَّذِينَ ءَامَنُوا إِذَا قِيلَ لَكُمْ تَفَسَّحُوا فِي الْمَجَالِسِ فَافْسَحُوا يَفْسَحِ اللَّهُ لَكُمْ وَإِذَا قِيلَ

أَنْشُرُوا فَأَنْشُرُوا يَرْفَعِ اللَّهُ الَّذِينَ ءَامَنُوا مِنْكُمْ وَالَّذِينَ أُوتُوا الْعِلْمَ دَرَجَاتٍ ۚ وَاللَّهُ بِمَا

تَعْمَلُونَ خَبِيرٌ (١١)

Artinya: “Hai orang-orang beriman apabila dikatakan kepadamu: “Berlapang-lapanglah dalam majlis”, maka lapangkanlah niscaya Allah akan memberi

kelapangan untukmu. Dan apabila dikatakan: “Berdirilah kamu”, maka berdirilah, niscaya Allah akan meninggikan orang-orang yang beriman di antaramu dan orang-orang yang diberi ilmu pengetahuan beberapa derajat. Dan Allah Maha Mengetahui apa yang kamu kerjakan”. (Q.S Al mujadalah ayat 11).

Tafsir Singkat (al-Jalalain): *“(Hai orang-orang yang beriman, apabila dikatakan kepada kalian, “Berlapang-lapanglah) berluas-luaslah (dalam majelis”) yaitu majelis tempat Nabi SAW. berada, dan majelis zikir sehingga orang-orang yang datang kepada kalian dapat tempat duduk. Menurut suatu qiraat lafal al-majaalis dibaca al-majlis dalam bentuk mufrad (maka lapangkanlah, niscaya Allah akan memberi kelapangan untuk kalian) di surga nanti. (Dan apabila dikatakan, “Berdirilah kalian”) untuk melakukan shalat dan hal-hal lainnya yang termasuk amal-amal kebaikan (maka berdirilah) menurut qiraat lainnya kedua-duanya dibaca fansyuzuu dengan memakai harakat damah pada huruf Syinnya (niscaya Allah akan meninggikan orang-orang yang beriman di antara kalian) karena ketaatannya dalam hal tersebut (dan) Dia meninggikan pula (orang-orang yang diberi ilmu pengetahuan beberapa derajat) di surga nanti. (Dan Allah Maha Mengetahui apa yang kalian kerjakan).”*

Selain ditingkatkan derajatnya oleh alloh SWT keutaman ilmu juga tergambar dalam surat pertama Al-Quran yang diterima Nabi Muhammad SAW yaitu surat Al-Alaq pada ayat 1 sampai dengan 5 sebagai berikut.

اقْرَأْ بِاسْمِ رَبِّكَ الَّذِي خَلَقَ (١) خَلَقَ الْإِنْسَانَ مِنْ عَلَقٍ (٢) اقْرَأْ وَرَبُّكَ الْأَكْرَمُ (٣) الَّذِي عَلَّمَ بِالْقَلَمِ (٤) عَلَّمَ الْإِنْسَانَ مَا لَمْ يَعْلَمْ (٥)

Artinya:

“[1] Bacalah, dengan (menyebut) nama Tuhanmu yang menciptakan.[2] Dia telah menciptakan manusia dari segumpal darah.[3] Bacalah, dan Tuhanmulah Yang Mahamulia.[4] yang mengajar (manusia) dengan pena. [5] Dia mengajarkan manusia apa yang tidak diketahuinya.”(QS.Al-Alaq 1:5)

Begitu pentingnya sebuah ilmu bagi manusia sehingga mereka yang memiliki ilmu akan diangkat derajatnya oleh Alloh SWT. Dalam mencari sebuah ilmu haruslah sering dilakukan evaluasi dalam sebuah pembelajaran, hal ini dikarenakan dengan sering melakukan evaluasi diri mereka akan mengetahui kekurangan terhadap pelajaran yang telah dihadapi. Hal ini selaras dengan ayat Al-Quran Q.S Al-Ankabut ayat 2 dan 3 sebagai berikut.

أَحْسِبِ النَّاسُ أَنْ يُنْزَكُوا أَنْ يَقُولُوا آمَنَّا وَهُمْ لَا يُفْتَنُونَ (٢) وَلَقَدْ فَتَنَّا الَّذِينَ مِنْ قَبْلِهِمْ
فَلَيَعْلَمَنَّ اللَّهُ الَّذِينَ صَدَقُوا وَلَيَعْلَمَنَّ الْكَاذِبِينَ (٣)

Artinya:

“[2] Apakah manusia mengira bahwa mereka dibiarkan hanya dengan mengatakan, "Kami telah beriman," dan mereka tidak diuji? [3] Dan sungguh, Kami telah menguji orang-orang sebelum mereka, maka Allah pasti mengetahui orang-orang yang benar dan pasti mengetahui orang-orang yang dusta.”(QS.Al-Ankabut 2:3).

Jika di kombinasikan pada surat (Al-Alaq Ayat 1:5) yang menerangkan tentang pentingnya membaca dalam mencari ilmu dengan surat (Al-Ankabut ayat 2:3) yang menerangkan sebuah evaluasi diri maka penulis berharap dari kombinasi

potongan ke dua surat tersebut, untuk kedepannya selain pandai untuk melakukan evaluasi terhadap diri sendiri, evaluasi juga penting pada sebuah proses dalam mencari ilmu. Dengan ditunjang dengan teknologi terbaru dan pengembangan aplikasi BEESMART sebagai objek penelitian akan bermanfaat khususnya pada memaksimalkan proses belajar.



BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Untuk menentukan tingkat kesulitan butir soal secara otomatis pada aplikasi BEESMART yaitu melakukan klasifikasi terhadap jawaban peserta dengan menjumlahkan butir soal berdasarkan kategori benar dan salah. Selanjutnya, jumlah kategori butir soal benar dan salah yang diketahui dihitung dengan melakukan klasifikasi terhadap jumlah jawaban benar siswa berdasarkan kemampuan siswa atas dan bawah. Kemudian jumlah jawaban butir soal peserta tes yang telah di klasifikasikan berdasarkan kemampuan atas dan bawah dibagi dengan jumlah peserta yang memiliki kemampuan atas dan bawah. Dari proses tersebut maka akan dihasilkan sebuah nilai tingkat kesulitan yang selanjutnya nilai tingkat kesulitan butir soal akan diklasifikasikan kembali berdasarkan kategori tingkat kesulitan butir soal baru. Kemudian hasil dari klasifikasi kategori tingkat kesulitan butir soal baru akan digunakan untuk memperbarui status kategori tingkat kesulitan yang lama. Selain itu, fitur lain yang dimiliki dari penerapan *Item Response Theory Model 2 Parameter Logistic* pada aplikasi BEESMART yaitu mampu memberikan sebuah informasi berupa kategori tingkat kesulitan soal yang diperoleh dari hasil tes kemampuan siswa dengan informasi kategori butir soal yaitu mudah, sedang, susah. Selain informasi berupa tingkat kesulitan soal, terdapat juga informasi berupa kategori tingkat kelayakan butir soal, dengan keterangan informasi butir soal berupa butir soal baik, butir soal cukup baik, butir soal kurang baik dan butir soal ditolak yang digambarkan dengan keterangan dan informasi dalam bentuk kurva sehingga mampu mengetahui dayabeda dan tingkat kesulitan butir soal guna untuk

meningkatkan kualitas butir soal dan mendapatkan informasi butir soal secara cepat.

5.2 Saran

Berdasarkan penerapan *computer adaptive* tes pada aplikasi *BEESMART* menggunakan metode *item response theory* model 2 parameter *logistic* dan aplikasi berbasis *web*, sehingga ada kesempatan bagi para peneliti untuk menerapkan *adaptive* tes dengan parameter 3 atau bahkan metode lain dan mengembangkan aplikasi *BEESMART* yang berbasis *mobile*.

DAFTAR PUSTAKA

- Anastasi, A & S. Urbina. 1997. *Psychological Testing*. Seventh Edition. Precentice Hall International Inc, New Jersey.
- Anggreyani, A. 2009. *Penerapan Teori Uji Klasik Dan Teori Respon Butir Dalam Mengevaluasi Butir Soal*. (Skripsi Jurusan Statistika Ipb)
- Baker, F.B. 2001. *The Basics Of Item Response Theory*. Eric Clearinghouse On Assessment And Evaluation.
- Calongesi, J.S. 1995. *Merancang Test Untuk Menilai Prestasi Siswa*. Bandung: Itb
- Crocker, L. & Algina, 1986. *Introduction to Classical and Modern Test Theory*. Holt, Rinehart and Winston, Inc. New York
- Erawan, Lalang Dan Santosa, Stefanus. 2010. *Computer Adaptive Test Dengan Pendekatan Item Response Theory Satu Parameter*. Journal Teknologi Informasi, Volume 5 Nomor 2, Oktober 2009, Issn 1414-9999.
- Hambleton R.K, & Swaminathan, H. 1991. *Item Response Theory, Principles And Applications*. Springer.
- Hulya Olmus, Ezgi Nazman Dan Semra Erbas. 2017. *An Evaluation Of The Two Parameters (2-Pl) Irt Models Through A Simulation Study*. Gu J Sci : 235-249.
- Mehrens, W.A., dan I.J Lehmann, *Meansurment and evaluation in Education and Psychology*, Second Edition, Holt, Rinehart and Winston, New York-Chicago-San Fransisco-Dallas-Montreal-London-Sydney, 1987.
- Nitko, A.J. 1996. *Educational Assessment Of Students, Secon Edition*. Merill An Imprint Of Prentice Hall Englewood Cliffs, Ohio.

Norman E.Gronlund, 2012, *Prinsip-Prinsip Dan Teknik Evaluasi Pembelajaran*, PT Remaja Rosdakarya, Bandung.

Rudner, Lawrence M. 1998. An On-Line, An On-Line, Interactive, Computer Adaptive Testing Tutorial, 11/98 Tutorial, [Http://Echo.Edres.Org:8080/Scripts/Cat/Catdemo.Htm](http://Echo.Edres.Org:8080/Scripts/Cat/Catdemo.Htm) (14 Januari 2018)

Wildan Setyo, P. 20014. *Mobile BEESMART Menggunakan Metode Item Response Theory Untuk Evaluasi Pembelajaran Bahasa Arab* (Skripsi Jurusan Teknik Informatika Uin Maulana Malik Ibrahim Malang).

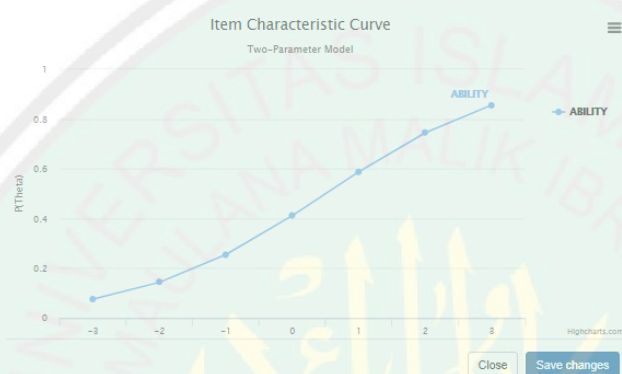
Yu, Chong Ho . 2010. *Simple Guide To The Item Response Theory(Irt) And Rasch Modeling*



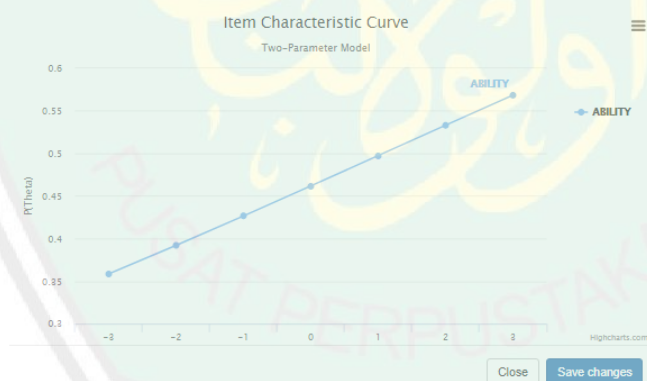
LAMPIRAN

Kurva Karakteristic Butir Soal Butir 1 – 20 Mata Kuliah Aahasa Arab

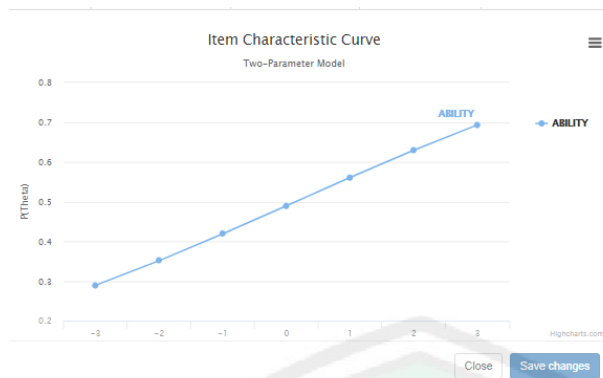
1. Kurva Karakteristic Butir Soal Butir 1



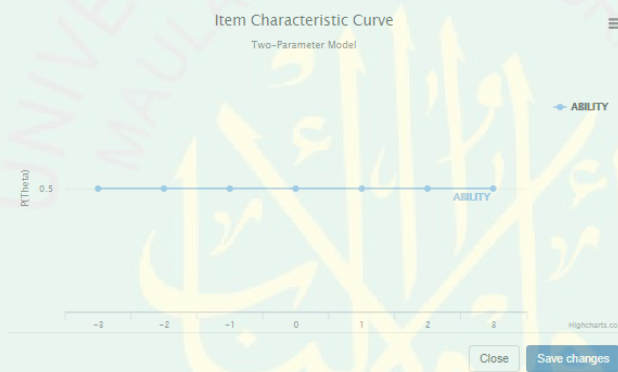
2. Kurva Karakteristic Butir Soal Butir 2



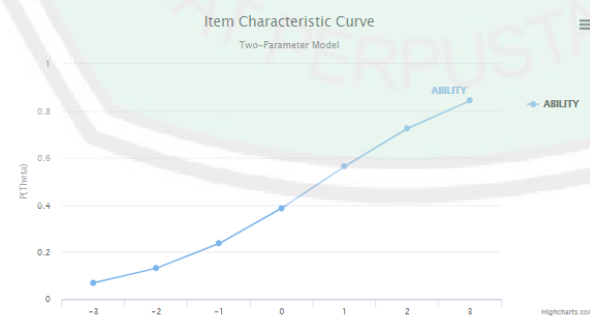
3. Kurva Karakteristic Butir Soal Butir 3



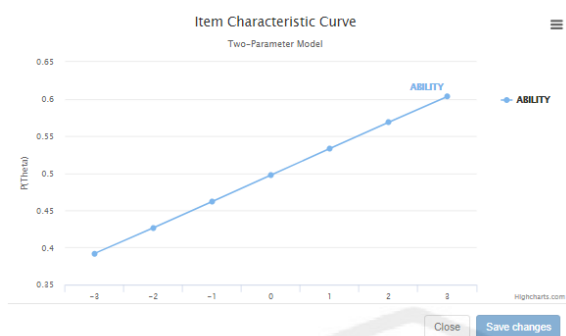
4. Kurva Karakteristic Butir Soal Butir 4



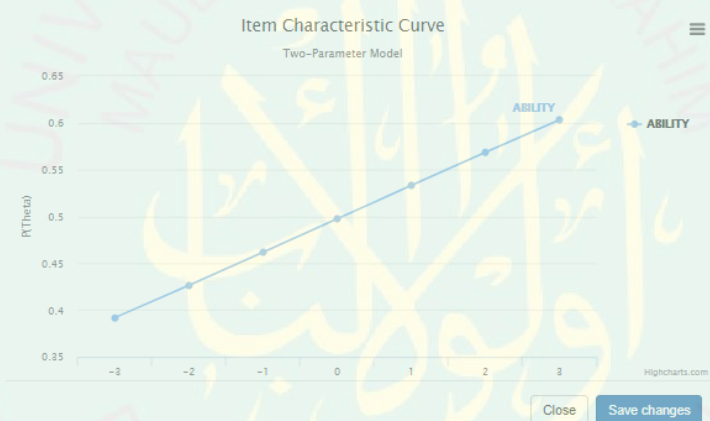
5. Kurva Karakteristic Butir Soal Butir 5



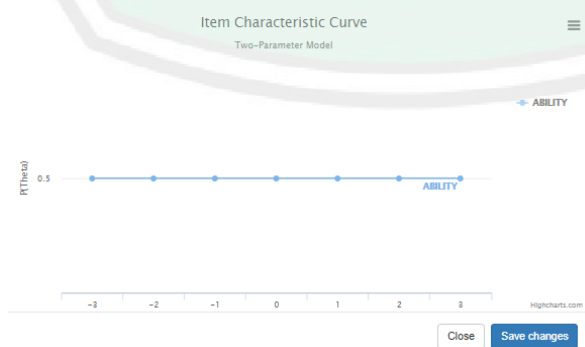
6. Kurva Karakteristic Butir Soal Butir 6



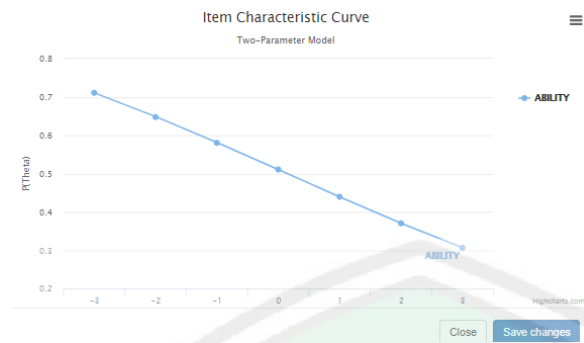
7. Kurva Karakteristic Butir Soal Butir 7



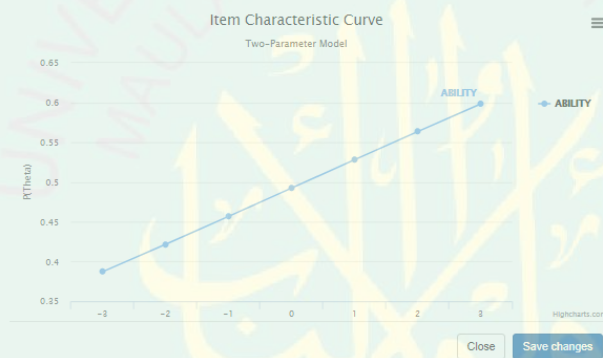
8. Kurva Karakteristic Butir Soal Butir 8



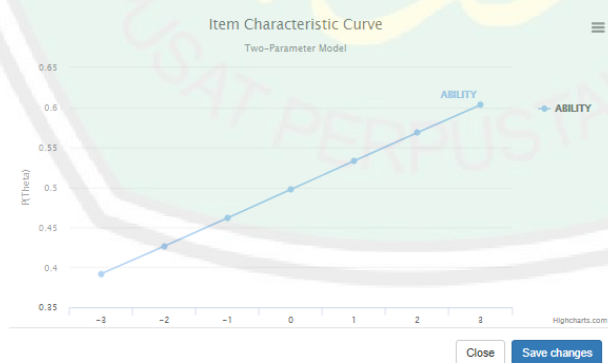
9. Kurva Karakteristic Butir Soal Butir 9



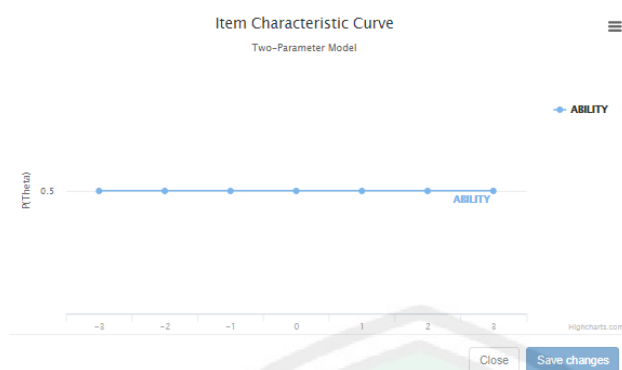
10. Kurva Karakteristic Butir Soal Butir 10



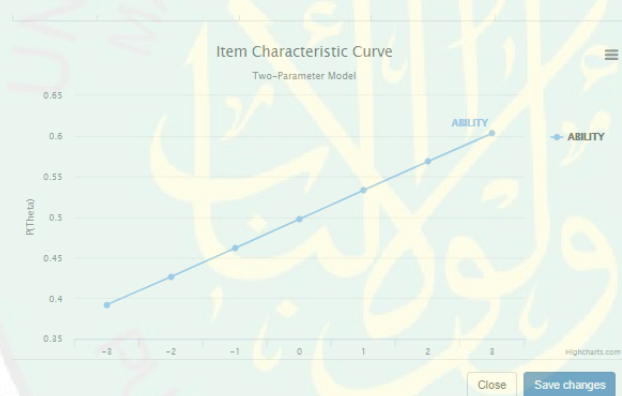
11. Kurva Karakteristic Butir Soal Butir 11



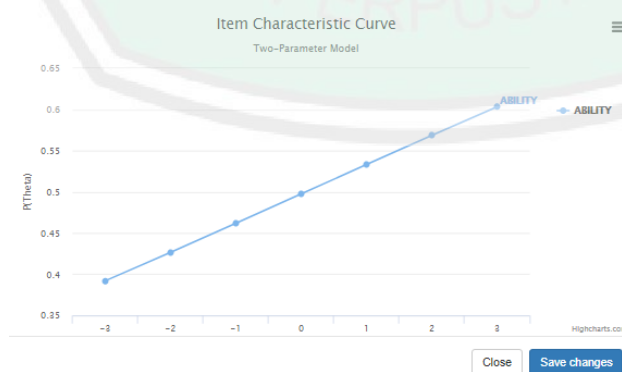
12. Kurva Karakteristic Butir Soal Butir 12



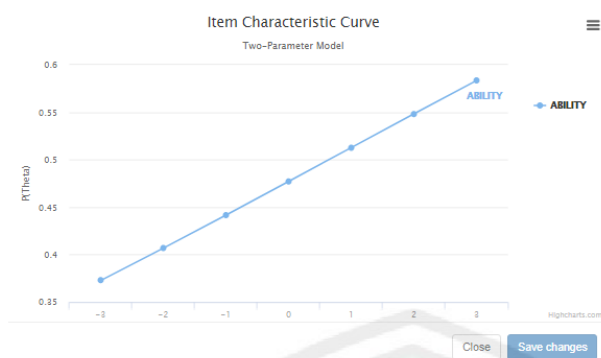
13. Kurva Karakteristic Butir Soal Butir 13



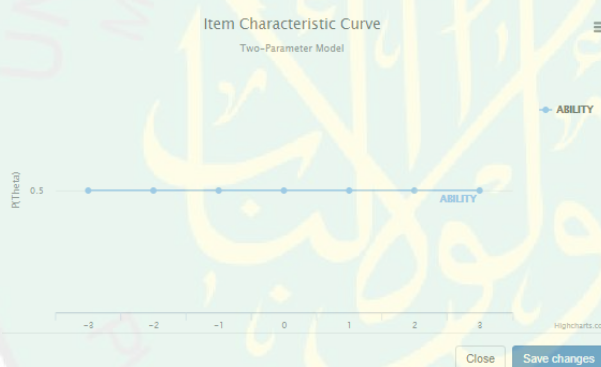
14. Kurva Karakteristic Butir Soal Butir 14



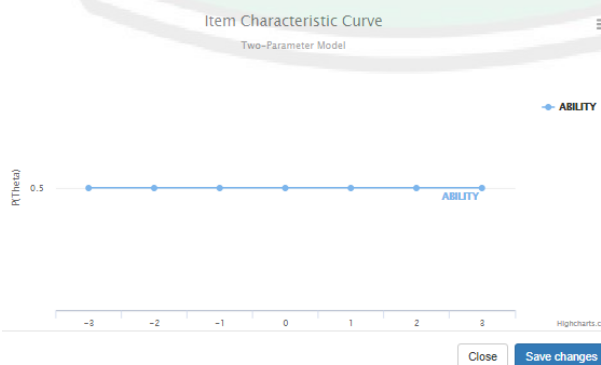
15. Kurva Karakteristic Butir Soal Butir 15



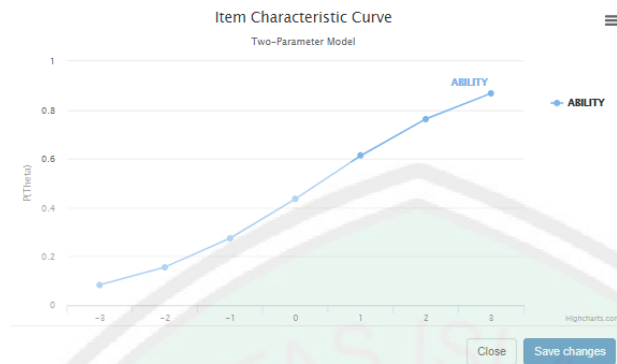
16. Kurva Karakteristic Butir Soal Butir 16



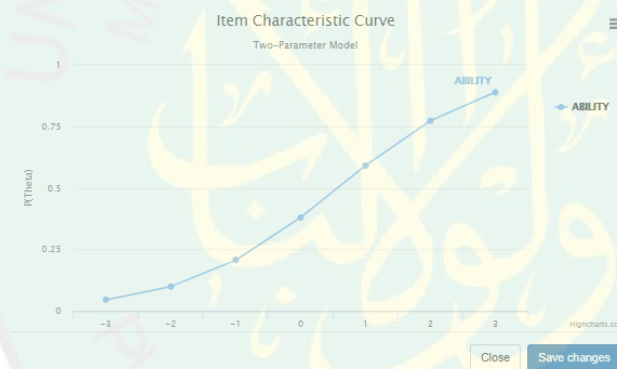
17. Kurva Karakteristic Butir Soal Butir 17



18. Kurva Karakteristic Butir Soal Butir 18



19. Kurva Karakteristic Butir Soal Butir 19



20. Kurva Karakteristic Butir Soal Butir 20

