

**PENGEMBANGAN *COMPUTERIZED ADAPTIVE TESTING*
(CAT) MENGGUNAKAN METODE
POHON SEGITIGA KEPUTUSAN**

Winarno

Sekolah Tinggi Agama Islam Negeri (STAIN) Salatiga
winarno5@yahoo.com

Abstrak

Penelitian ini bertujuan menghasilkan CAT menggunakan metode pohon segitiga keputusan dalam prosedur pemilihan item dan mengetahui kemampuan CAT dalam mengestimasi kemampuan peserta tes dengan tepat. Penelitian ini menggunakan *Research and Development* (R&D). Pengambilan data dengan observasi, dokumentasi, dan angket. Analisis data yang digunakan adalah teknik analisis deskriptif evaluatif dan teknik analisis deskriptif kuantitatif. Hasil penelitian adalah (1) CAT yang dikembangkan berdasar kebutuhan pemakai yaitu: berbasis internet, memiliki sistem keamanan, dan mudah diakses, (2) CAT dapat mengenali tiga pengguna, yaitu: administrator, guru, dan siswa, (3) CAT mampu memberikan butir-butir yang bersifat adaptif berdasarkan respon jawaban peserta tes. Secara keseluruhan kinerja CAT mampu melaksanakan tugas dengan baik untuk memilih butir tes dan mengukur kemampuan peserta tes dengan akurat dan tepat dilihat dari nilai korelasi antara hasil estimasi kemampuan (θ) dengan nilai ulangan murni (NUM) di sekolah siswa cukup tinggi yakni 0,67.

Kata kunci: *metode pohon segitiga keputusan, metode maximum likelihood*

DEVELOPING COMPUTERIZED ADAPTIVE TESTING (CAT) BY USING THE TRIANGLE DECISION TREE METHOD

Winarno

Sekolah Tinggi Agama Islam Negeri (STAIN) Salatiga
winarno5@yahoo.com

Abstract

This research aims at producing a CAT software that uses the triangle decision tree method in the test item selection procedure and detecting the CAT ability in estimating the test-takers' ability accurately and correctly. This research used the research and development approach (R&D). The data were collected through observation, documentation, inquiry, and the data were analyzed descriptively and quantitatively. The findings are as follows. (1) The CAT developed is based on users' need, web-based, user-friendly, interactive, highly secured, and easily accessible. (2) The CAT can recognize three different users: school administrators, teachers, and students. (3) The CAT software is able to provide adaptive items according to students' answers. Overall, the CAT software can perform the tasks well to select test items and to measure test-takers' ability. It can be seen from correlation value between ability estimation result (θ) in CAT with *nilai ulangan murni* (NUM) students at school which is high enough that is 0,67.

Keywords: *the triangle decision tree method, maximum likelihood method*

Pendahuluan

Evaluasi merupakan salah satu rangkaian kegiatan dalam meningkatkan kualitas, kinerja atau produktivitas suatu lembaga dalam melaksanakan programnya. Tujuan evaluasi adalah untuk melihat dan mengetahui proses yang terjadi dalam proses pembelajaran. Melalui evaluasi akan diperoleh informasi tentang apa yang telah dicapai dan mana yang belum (Mardapi, 2004: 19). Evaluasi memberikan informasi bagi kelas dan pendidik untuk meningkatkan kualitas proses belajar mengajar. Evaluasi sebagai komponen pengajaran adalah proses untuk mengetahui keberhasilan program pengajaran dan merupakan proses penilaian yang bertujuan untuk mengetahui kesukaran-kesukaran yang melekat pada proses belajar (Murshel, 1954: 373).

Evaluasi dalam pendidikan dilaksanakan untuk memperoleh informasi tentang aspek yang berkaitan dengan pendidikan. Menurut Gronlund (1976: 8), evaluasi dalam pendidikan memiliki tujuan : a) untuk memberikan klarifikasi tentang sifat hasil pembelajaran yang telah dilaksanakan, b) memberikan informasi tentang ketercapaian tujuan jangka pendek yang telah dilaksanakan, c) memberikan masukan untuk kemajuan pembelajaran, d) memberikan informasi tentang kesulitan dalam pembelajaran dan untuk memilih pengalaman pembelajaran di masa yang akan datang.

Informasi evaluasi dapat digunakan untuk membantu memutuskan kesesuaian dan keberlangsungan dari tujuan pembelajaran, kegunaan materi pembelajaran, dan untuk mengetahui tingkat efisiensi dan efektifitas dari strategi pengajaran (metode dan teknik belajar-mengajar) yang digunakan. Evaluasi memiliki fungsi untuk membantu guru dalam hal-hal: a) penempatan siswa dalam kelompok-kelompok tertentu, b) perbaikan metode mengajar, c) mengetahui kesiapan siswa (sikap, mental, material), d) memberikan bimbingan dan seleksi dalam rangka menentukan jenis jurusan maupun kenaikan tingkat (Gronlund, 1976: 16). Dalam evaluasi pendidikan, diperlukan alat (instrumen), alat yang digunakan untuk melakukan evaluasi, salah satunya adalah tes.

Selama ini, sebagian besar alat untuk mengukur tes menggunakan kertas dan pensil (*paper and pencils test*). Perkembangan terkini dalam usaha peningkatan pelaksanaan tes dengan kehadiran teknologi komputer telah mulai dipergunakan untuk kemajuan pengujian (Hambleton, Swaminathan, dan Rogers, 1991: 146). Komputerisasi penilaian individu lebih efisien dan akurat daripada penilaian menggunakan kertas dan pensil (*pencil and paper test*) (Wainer, 1990: 273). Salah satu *prototype* komputerisasi penilaian individu yang berkembang saat ini adalah *Computerized Adaptive Testing* (CAT). CAT adalah suatu metode pengujian atau evaluasi dengan menggunakan teknologi informasi yang bersifat adaptif. Adaptif berarti bahwa pemberian soal ujian berikutnya tergantung pada perilaku peserta ujian dalam menjawab soal sebelumnya sehingga ujian yang diberikan untuk setiap peserta dapat bersifat unik berdasarkan tingkat kemampuan masing-masing peserta.

Kelebihan-kelebihan yang ditawarkan oleh CAT antara lain: (1) CAT lebih efisien dan akurat dalam mengukur kemampuan peserta tes (Weiss, 2004: 2). CAT tidak memerlukan lembar jawaban karena skor dapat segera diketahui oleh peserta tes begitu tes telah dinyatakan selesai, (2) soal yang diberikan memiliki level kesukaran sesuai dengan kemampuan siswa, tidak terlalu susah ataupun terlalu mudah, (3) penilaian dapat dilakukan dengan segera sehingga dapat memberikan umpan balik yang cepat kepada siswa, (4) keamanan ujian dapat ditingkatkan. Rangkaian soal yang diberikan akan berbeda untuk setiap siswa sehingga soal yang akan muncul selanjutnya tidak dapat ditebak. Selain itu, bila jumlah soal banyak, kemungkinan munculnya soal yang sama lebih dari satu kali sangat kecil sehingga kemungkinan siswa untuk menghafal soal menjadi sangat kecil. Kerahasiaan soal pun dapat terjaga, karena soal tersimpan dalam suatu basis data dan hanya pembuat CAT yang membuat soal tersebut yang dapat meng^updatenya, dan (5) ujian dapat dipresentasikan melalui teks, grafik, audio, dan bahkan video klip.

CAT memerlukan: (1) bank soal, (2) prosedur pemilihan item awal, (3) prosedur pemilihan item selama pelaksanaan tes, (4) prosedur untuk mengakhiri tes, dan (5) estimasi kemampuan peserta tes (Masters & Keeves, 1999: 130). Dalam prosedur pemilihan item awal diberikan item

tes dengan tingkat kesukaran yang sedang. Prosedur pemilihan item tes selama pelaksanaan tes dalam CAT berdasarkan pada pola jawaban peserta tes yang akan dijadikan aturan untuk menentukan item tes berikutnya.

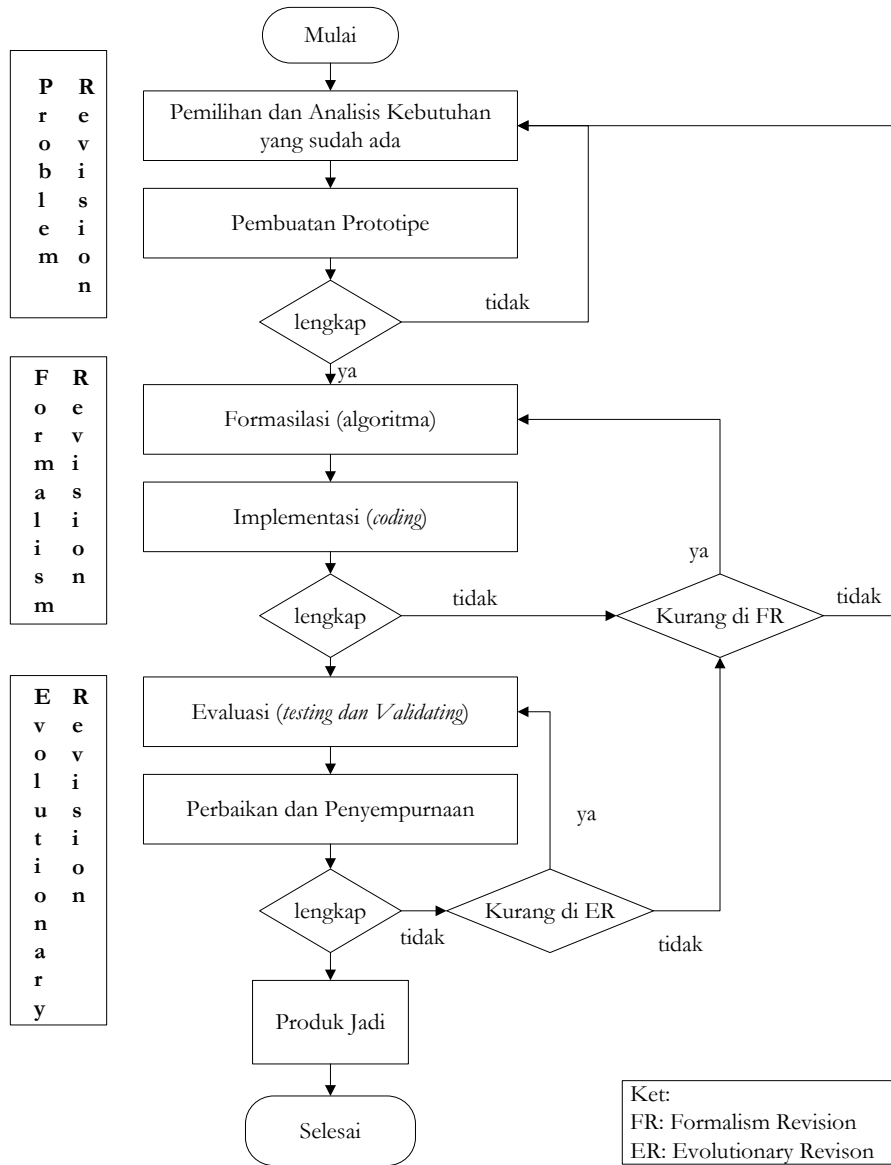
Salah satu metode pemilihan item selama pelaksanaan tes adalah menggunakan pohon segitiga keputusan. pohon segitiga keputusan adalah model keputusan yang berbentuk grafik. Sebuah simpul dalam pohon segitiga keputusan menunjukkan parameter-parameter tes yang berisi tiga parameter IRT yaitu tingkat kesukaran, daya beda, dan tingkat menebak. Setiap simpul hanya memiliki dua cabang ranting dan setiap cabang ranting yang keluar dari simpul ada dua arah yaitu cabang ke arah kiri dan cabang ke arah kanan. Arah cabang akan ke kanan jika peserta tes menjawab pertanyaan dengan benar dan arah cabang akan ke kiri jika peserta tes menjawab pertanyaan item yang salah (Phankokkruad, 2008: 656). Prosedur mengakhiri tes diberikan agar tes tidak terlalu panjang dan estimasi kemampuan peserta tes menggunakan metode *maximum likelihood* (MLE).

Metode Penelitian

Jenis penelitian dalam ini menggunakan *Research and Development* (R&D). Dalam pengembangan *software* CAT ada dua tahap pelaksanaan yang dilakukan yakni tahap pertama adalah tahap pengembangan produk dan tahap kedua adalah tahap implementasi produk .

Tahap Pengembangan Produk

Pada tahap pertama dalam pengembangan produk, langkah yang diambil mengikuti langkah-langkah yang dikemukakan oleh Kendal dan Kendal serta Pressman dan telah dilengkapi oleh Rolston (1988: 138) yang meliputi: 1) pemilihan dan analisis kebutuhan yang sudah ada, 2) pembuatan prototipe, 3) formasilasi (penggunaan metode), 4) implementasi (*coding*), 5) evaluasi (*testing dan validating*), dan 6) perbaikan dan penyempurnaan seperti Gambar 1 berikut.



Gambar 1. Langkah-langkah pengembangan CAT

Berdasar pada Gambar 1 di atas, langkah pertama dan kedua dinamakan *problem revision* kemudian langkah ketiga dan keempat dinamakan *formalism revision* sedangkan langkah kelima dan keenam dinamakan *evolutionari revision*.

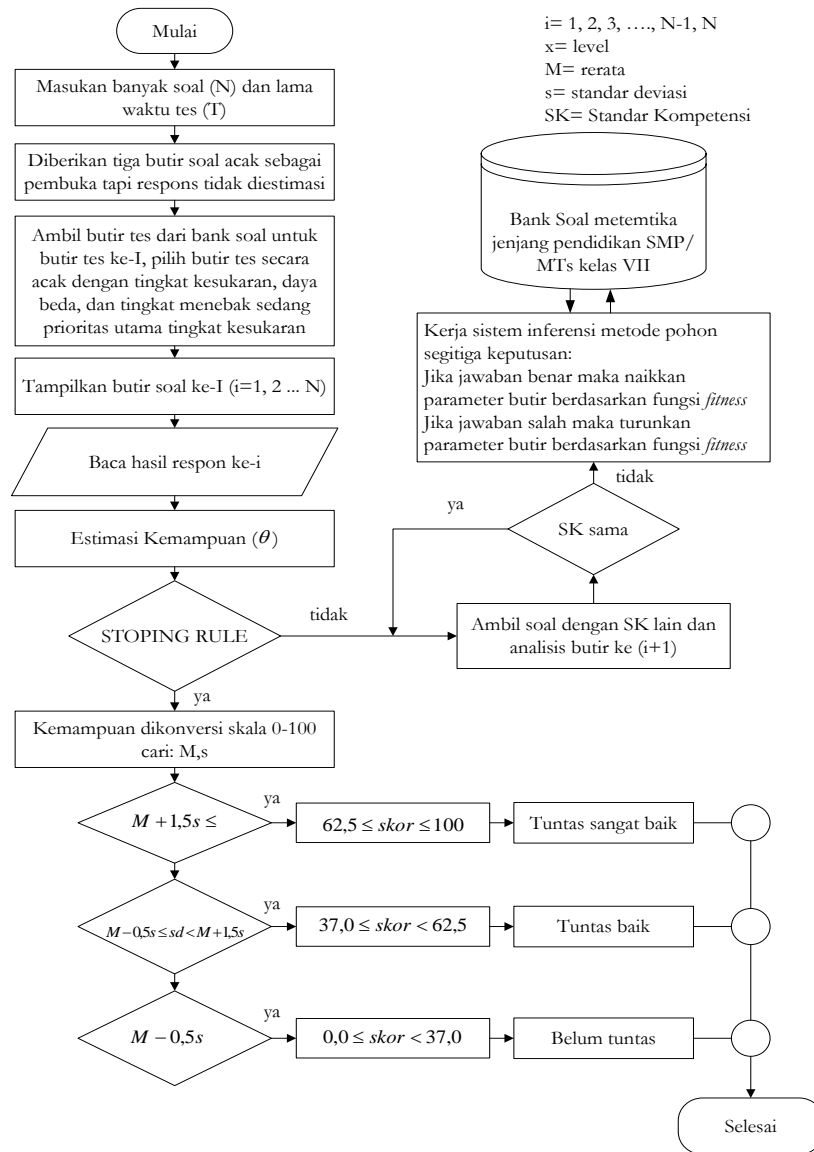
a. Langkah *problem revision*

Pada langkah *problem revision* dilakukan pemilihan dan analisis kebutuhan sistem yang sudah ada untuk pengumpulan informasi yang berfungsi untuk *need assesment* yang sudah ada sebagai desain penyusunan model. Berdasarkan informasi yang terkumpul dibuat prototipe perangkat lunak. Salah satu *need assesment* adalah prosedur kerja CAT menggunakan pohon segitiga keputusan. Dalam penelitian disertasi ini, bank soal yang digunakan adalah bank soal matematika jenjang pendidikan SMP/MTs kelas VII. Jumlah soal ada 193 butir soal sehingga cocok diaplikasikan dalam CAT metode pohon keputusan untuk level 18, dengan ketentuan 171 butir soal Kategori baik masuk ke dalam sistem basis data utama yang akan dimunculkan oleh komputer dan hasil respons akan diestimasi menjadi kemampuan peserta tes, sedangkan 22 butir soal kategori kurang baik akan menjadi soal pembuka dalam CAT dan hasil respons tidak diestimasi sebagai kemampuan peserta tes. Prosedur kerja CAT menggunakan pohon segitiga keputusan lebih detail disajikan pada Gambar 2.

Penjelasan sistem kerja CAT menggunakan pohon segitiga keputusan sebagai berikut.

1) *Starting rules*

Pemilihan butir awal menggunakan asumsi bahwa peserta tes tidak diketahui kemampuan awal sehingga diambil butir soal dengan tingkat kesukaran yang sedang agar peserta tes dengan kemampuan ekstrim kurang pandai tidak terlalu lama mencapai *stopping rule* dan peserta tes dengan kemampuan ekstrim pandai terlalu singkat mencapai *stopping rule*.



Gambar 2. Prosedur kerja CAT menggunakan pohon segitiga keputusan

2) Pemilihan butir selama pelaksanaan tes

Pemilihan butir selama pelaksanaan tes menggunakan metode pohon segitiga keputusan. Dalam segitiga keputusan terdapat segitiga utama yakni berisi semua kemampuan dalam matematika, dalam segitiga utama terdapat sub-sub segitiga keputusan yang berisi standar kompetensi (SK), dan Kompetensi dasar (KD) dari matematika sehingga ada segitiga-segitiga kecil yang berisi SK tentang bilangan, SK tentang aljabar, dan SK tentang geometri .

Dalam CAT menggunakan metode pohon segitiga keputusan, soal berikutnya diberikan dengan kategori tingkat kesukaran, daya beda, dan tingkat menebak akan naik berdasar fungsi *fitness* jika peserta tes menjawab soal dengan benar dan soal berikutnya diberikan dengan kategori tingkat kesukaran, daya beda, dan tingkat menebak akan turun jika peserta tes menjawab soal dengan salah berdasar fungsi *fitness*

3) *Stopping rules*

Kriteria *stopping rules* yang diambil yakni: (a) soal habis jika respons jawaban peserta tes berpola, (b) kemampuan menuju nilai Θ tertentu jika respons jawaban peserta tes tidak berpola, dan (c) selisih error adalah 0,01

b. Estimasi kemampuan peserta tes.

Estimasi kemampuan peserta tes menggunakan metode *maximum likelihood* (MLE) setelah diketahui hasil benar atau salah dari respons jawaban peserta

Proses kembali ke pemilihan dan analisis kebutuhan jika dalam pengembangan prototipe ada kekurangan informasi. Langkah ini dinamakan *problem revision*. Pada langkah ini dilakukan terus menerus untuk memperoleh langkah yang representatif. Langkah akan berlanjut ke langkah berikutnya apabila ruang lingkup permasalahan yang diselesaikan telah terpenuhi.

c. Langkah *Formalism Revision*

Langkah kedua dalam perancangan logaritma dilakukan pembuatan sistematika kerja program perangkat lunak yang berdasarkan langkah pertama. Dengan langkah pada algoritma kemudian menerjemahkan

algoritma ke dalam kode program. Pada langkah kedua ini akan kembali ke langkah perancangan algoritma apabila terdapat kode program yang tidak sesuai dengan algoritma. Langkah kedua ini dinamakan *formalism revision*. Proses ini juga bisa menuju ke langkah pertama bila terdapat informasi yang belum lengkap dan kurang sesuai dengan langkah pada langkah pertama. Proses pada langkah kedua ini akan menuju ke langkah ketiga bila target telah terpenuhi yaitu mendapatkan sebuah program yang mampu digunakan untuk menyelesaikan masalah.

d. Langkah *Evolutionari Revision*

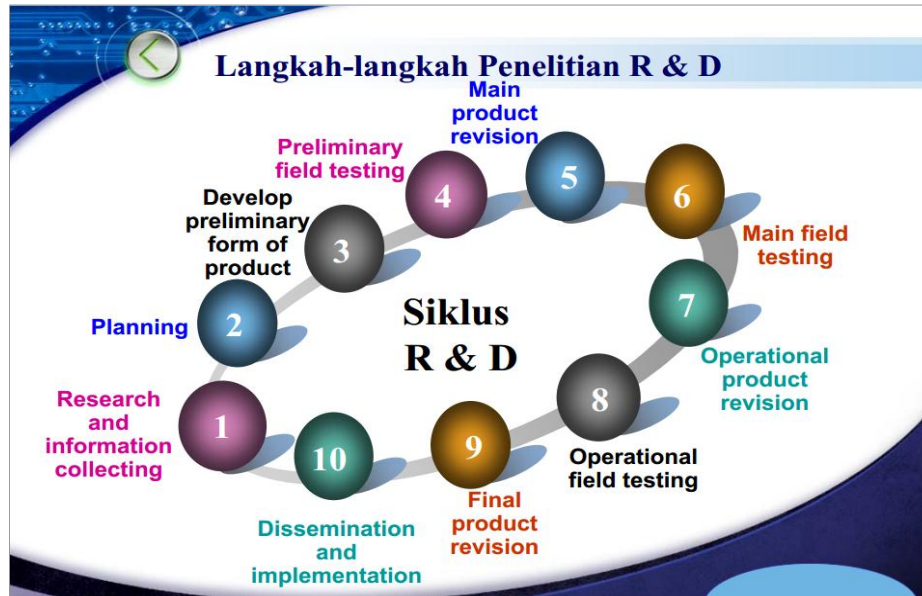
Langkah ketiga pengujian perangkat lunak (*debugging*) yaitu langkah untuk menemukan kesalahan yang mungkin terjadi. Ada tiga kesalahan yang mungkin terjadi yaitu: (1) *syntax error* (kesalahan kalimat), (2). *run time error* (kesalahan saat dijalankan), dan (3). *logic error* (kesalahan fungsi dan hasil dari penalaran logika). Dari langkah ini dijadikan dasar proses perbaikan dan penyempurnaan program. Proses ini akan kembali ke pengujian program jika masih ada kesalahan yang menyebabkan program belum berfungsi seperti yang diharapkan. Langkah ketiga ini dinamakan *evolutionari revision*. Proses akan kembali ke langkah kedua (*Formalism revision*) jika ada kesalahan yang disebabkan oleh algoritma dan penulisan kode program yang belum sesuai dengan langkah ke dua. Atau bahkan akan ke langkah pertama (*problem revision*) apabila ada kesalahan algoritma dan kode program yang kurang sesuai yang disebabkan adanya algoritma dan kode program yang belum sesuai pada langkah pertama.

Langkah-langkah ini adalah berbentuk siklus hidup untuk mengembangkan CAT. Dalam siklus-siklus ini mengalami proses berulang jika pada langkah tertentu ada kesalahan, proses akan berulang pada bagian yang ditemukan kesalahan. Siklus akan bergerak terus menerus sehingga diperoleh perangkat lunak yang secara operasional dapat berfungsi sesuai dengan tujuan yang telah ditentukan.

e. Tahap Implementasi Produk

Tahap kedua adalah tahap implementasi produk, dalam tahap ini dilakukan implementasi CAT dalam situasi kelas yang sesungguhnya.

Proses ini mengikuti langkah dari Borg & Gall (2003: 775) seperti Gambar 3 berikut.



Gambar 3. Langkah –langkah penelitian R & D

Tempat penelitian (pengembangan) dilakukan di laboratorium komputer Sekolah Tinggi Agama Islam Negeri (STAIN) Salatiga dalam jangka waktu penelitian selama 1 tahun pada Oktober 2010 sd September 2011.

Prosedur pengembangan dalam penelitian ini dilakukan dalam jangka waktu tertentu. Tahap-tahap penelitian yang dilakukan dipilih essential dan harus melewati sebuah rancangan produk yaitu: (1) analisis dan identifikasi kebutuhan, (2) perancangan desain sistem, (3) pembuatan produk (*coding*), (4) pengujian *feasibilitas* produk, (5) uji coba produk, (6) implementasi produk, dan (7) revisi produk.

Pengujian CAT ini menggunakan *test and debugging*. Program yang telah terkemas ke dalam satu kesatuan produk, dioperasikan dan diuji keberhasilannya dalam melakukan tugas sesuai input yang diberikan pengguna. Pada tahap uji program ini akan dicari tiga tipe kesalahan yang dikejar yaitu : *syntax*, *run-time (semantik)*, dan *logic*. Pada *syntax error*, semua *syntax error* harus ditemukan dan dibetulkan sebelum sebuah program akan dieksekusi. Untuk *Run-time error* dilacak dari algoritma yang digunakan untuk membuat program. Kesalahan tipe ketiga yang sangat sulit ditemukan adalah kesalahan logis (*error logic*) karena dalam hal ini program berhasil dijalankan, namun hasil keluaran tidak/belum sesuai dengan yang diinginkan.

Sebelum dilakukan uji coba, *software* CAT yang telah dibuat dilakukan kegiatan validasi pakar dengan teknik FGD (*Focus Group Discussion*) oleh para pakar diantaranya: pakar matematika, pakar bahasa pemrograman, pakar pengukuran, dan pakar CAT. Kemudian model diuji oleh beberapa orang pemakai diantaranya siswa SMP/MTs kelas VIII dan guru pada jenjang pendidikan SMP/MTs yang mengajar kelas VII. Sekolah tempat implementasi CAT dipilih dengan teknik *purposive sampling*, yakni sekolah yang memiliki laboratorium komputer dan dilengkapi jaringan LAN (*Local Area Network*). Tempat implementasi dilakukan di SMP N 1 kota Salatiga yang memiliki laboratorium komputer dengan dilengkapi jaringan LAN.

Proses pengujian dilakukan menggunakan bantuan seperangkat unit komputer dan dilaksanakan di laboratorium komputer. Dalam pelaksanaannya setiap peserta tes akan diminta oleh komputer beberapa informasi antara lain: nama login, email, nama depan, nama belakang, tingkat pengguna, dan password. Kemudian peserta membuat *Account* yang berisi *user* dan *password*. Setelah peserta tes *login* maka komputer akan menampilkan soal yang diambil dari bank soal dalam CAT yaitu bank soal matematika jenjang pendidikan SMP/MTs kelas VII yang dikembangkan oleh peneliti dan berkolaborasi dengan MGMP matematika kota Salatiga.

Hasil respons peserta tes diberi nilai 1 jika jawaban benar dan diberi nilai 0 jika jawaban salah.. Hasil kemampuan (θ) kemudian dikonversi ke dalam skor dengan rentang 0 - 100. Perhitungan konversi skor seperti pada Tabel 1 di bawah ini.

Tabel 1. Konversi Kemampuan (θ) menjadi skor (y)

Jenis	Ketuntasan belajar		
	Belum tuntas		Belum tuntas
Kemampuan (θ)	$-4,00 \leq \theta < -1,00$	Kemampuan (θ)	$-4,00 \leq \theta < -1,00$
Penilaian Acuan Normal (x)	$0,0 \leq x < M-0,5s$	Penilaian Acuan Normal (x)	$0,0 \leq x < M-0,5s$
Skor skala 0 – 100	$0,0 \leq y < 37,0$	Skor skala 0 – 100	$0,0 \leq y < 37,0$

Keterangan: N = Nilai
s = Simpangan baku

M = Rerata
 $y = 12,5 \theta + 50$

Subyek yang digunakan dalam penelitian untuk untuk uji coba adalah: (a) 28 guru matematika untuk proses penggalian informasi dan identifikasi kebutuhan sistem yang diperlukan dalam CAT dan juga untuk memperoleh informasi mengenai pemilihan materi tes uji coba; (b) 51 siswa kelas VIII SMP Negeri 1 kota Salatiga dalam rangka uji fungsionalitas dan unjuk kerja program CAT; (c) pakar IT untuk mengetahui kelayakan CAT ini.

Teknik analisis data yang digunakan adalah teknik analisis deskriptif evaluatif dan teknik analisis deskriptif kuantitatif. Kedua teknik ini digunakan karena dalam penelitian ini tidak melakukan pengujian hipotesis. Penelitian ini akan menguji kelayakan produk yang digunakan untuk mengevaluasi kemampuan peserta menggunakan CAT yang menggunakan metode pohon segitiga keputusan.

Teknik analisis deskriptif evaluatif dilakukan untuk menentukan kelayakan, kemampuan dan efektifitas kerja produk dalam mengukur kemampuan peserta tes. Teknik analisis deskriptif kuantitatif dilakukan untuk mengetahui gambaran tingkat kemampuan peserta tes yang diukur dengan tes terkomputerisasi menggunakan metode pohon segitiga keputusan. Data kuantitatif yang diperoleh dari hasil penilaian pakar, perorangan, kelompok kecil, dan kelompok besar dalam kegiatan uji coba

dengan subyek uji coba guru dianalisis dengan menggunakan konsep reliabilitas antar responden, yakni generalizability coefisien menggunakan program GENOVA versi 3.1.

Hasil Penelitian

Hasil penelitian berupa: (1) hasil uji coba pakar, (2) hasil software CAT, dan (3) hasil uji coba siswa.

Hasil uji coba pakar

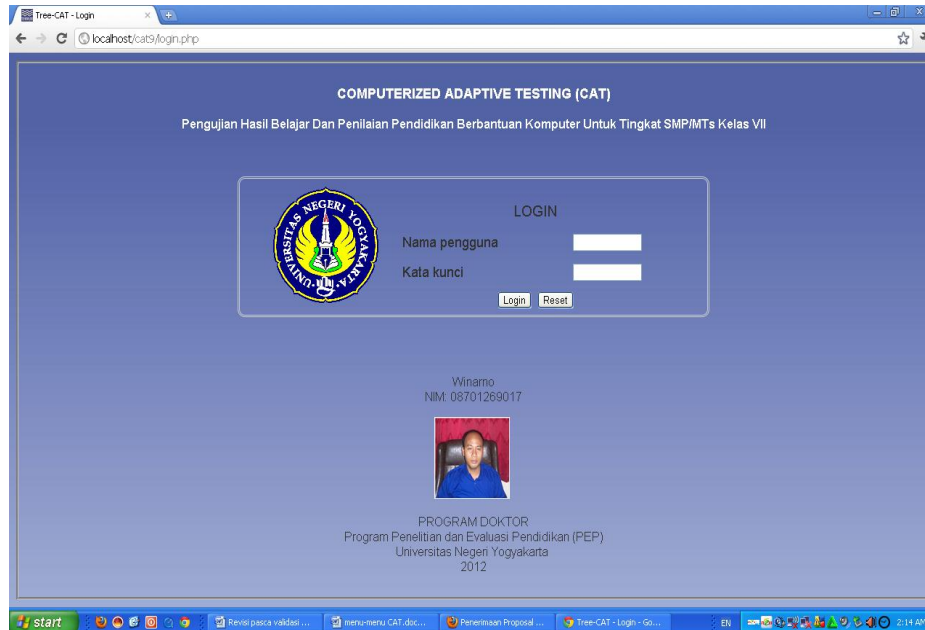
Hasil validasi pakar penelitian disajikan dalam Tabel 2 berikut.

Tabel 2. Hasil Validasi Pakar Penelitian

No	Responden	Aspek	Grand mean	$\hat{p}_{p^*}^2$	Ket
1	Pakar	Performasi Penggunaan	4,3	0,66	Tinggi
		Performansi Tampilan	3,7	0,46	Cukup tinggi
		Relevansi Materi Tes	4,4	0,78	Tinggi
		Kemanfaatan	4,2	0,75	Tinggi
2	Perorangan	Performasi Penggunaan	4,5	0,86	Tinggi
		Performansi Tampilan	4,3	0,46	Cukup tinggi
		Relevansi Materi Tes	4,2	0,88	Tinggi
		Kemanfaatan	4,5	0,75	Tinggi
3	Kelompok kecil	Performasi Penggunaan	3,9	0,78	Tinggi
		Performansi Tampilan	3,5	0,7	Tinggi
		Relevansi Materi Tes	3,6	0,73	Tinggi
		Kemanfaatan	4,1	0,82	Tinggi
4	Kelompok besar	Performasi Penggunaan	4,4	0,70	Tinggi
		Performansi Tampilan	4,1	0,74	Tinggi
		Relevansi Materi Tes	3,9	0,88	Tinggi
		Kemanfaatan	4,1	0,60	Cukup tinggi

Berdasar Tabel 2 di atas menunjukkan bahwa pesponden memiliki konsistensi yang cukup tinggi dan tinggi dalam memvalidasi CAT. Nilai grand mean tertinggi pada responden perorangan untuk aspek performansi penggunaan dan kemanfaatan yakni 4,5. Nilai \hat{p}_{p*}^2 tertinggi pada perorangan dan kelompok besar untuk aspek relevansi materi tes yakni sebesar 0,88.

Hasil Software CAT



Gambar 4. Tampilan awal CAT

Berdasarkan Gambar 4 di atas, halaman awal akan tampil sesaat setelah program CAT dipanggil (dijalankan). Desain tampilan awal terdiri dari: Teks pertama berbunyi “*COMPUTERIZED ADAPTIVE TESTING (CAT)*” yang diketik dengan huruf kapital semua dan berwarna putih. Tes kedua berbunyi “Pengujian Hasil Belajar dan Penilaian Pendidikan Berbantuan Komputer untuk mata pelajaran Matematika Tingkat

pendidikan SMP/MTs kelas VII” juga berwarna putih. Teks ketiga logo UNY, login, dan program Doktor Penelitian dan Evaluasi Pendidikan (PEP) Universitas Negeri Yogyakarta (UNY) Tahun 2012. Dibelakang ketiga teks terdapat warna latar belakang biru.

Tabel 3. Tampilan hasil rekam pembelajaran (RMP)

Soal Ke-	Id	No Pertanyaan	Benar/Salah	Waktu (detik)	TK	DB	TM	SK	KD	Indikator
1	36	12	Salah	13	-1,807	0,676	0,247	1	1	1
2	163	139	Salah	6	-1,815	1,121	0,236	2	2	3
3	98	74	Salah	7	-1,82	0,357	0,304	3	4	2
4	53	29	Benar	7	-1,894	0,562	0,239	4	2	2
5	71	47	Benar	7	-1,591	1,496	0,201	5	2	3
6	177	153	Benar	7	-1,335	1,143	0,248	6	3	2
7	83	59	Salah	5	-1,179	1,416	0,169	1	2	2
8	72	48	Salah	7	-1,179	1,09	0,269	2	4	5
9	41	17	Benar	7	-1,179	1,125	0,243	3	3	2
10	87	63	Benar	7	-1,078	1,101	0,236	4	1	2
11	85	61	Benar	7	-0,86	0,935	0,252	5	2	2
12	49	125	Benar	6	-0,745	0,976	0,272	6	3	2
13	125	101	Salah	8	-0,49	0,904	0,182	1	1	2
14	115	91	Benar	6	-0,51	0,845	0,247	2	3	3
15	207	183	Benar	6	-0,398	0,935	0,203	3	1	3
16	210	186	Benar	7	-0,089	1,112	0,285	4	5	1
17	124	100	Salah	13	0,1	0,782	0,206	5	1	1
18	42	18	Benar	5	0,066	5,848	0,444	6	2	4

Pada Tabel 3 di atas, tampilan hasil rekam pembelajaran terdiri atas soal ke-, ID, nomor pertanyaan, keterangan benar/salah, waktu, tingkat kesukaran, daya beda, tingkat menebak, standar kompetensi (SK), kompetensi dasar (KD), indikator, dan theta. Hasil rekam medik

pembelajaran (RMP) yang berisi hasil respons (benar atau salah) akan digunakan untuk mendiagnosis kemampuan siswa. Siswa yang paham materi pelajaran matematika maka mereka menjawab benar dan waktu yang singkat tetapi siswa yang tidak paham materi pelajaran matematika cenderung mengerjakan soal salah dan waktu yang cepat.

Dalam mendiagnosis kemampuan matematika maka dilihat dari respon benar-salah pada SK dan KD. Apabila siswa menjawab salah pada SK dan KD tertentu maka siswa tersebut belum menguasai dan apabila siswa menjawab benar pada SK dan KD tertentu maka siswa tersebut sudah paham.

Hasil Pengujian Siswa

Pengujian *beta* digunakan untuk mengetahui kemampuan kerja program CAT dalam memprediksi kemampuan siswa melalui butir-butir tes yang dikerjakan. Dalam hal ini program telah diisi basis data bank soal yang dilengkapi dengan beberapa identitas yakni: (1) nomor soal, (2) id soal, (3) nomor pertanyaan, (4) tingkat kesukaran (TK), (5) daya beda (DB), (6) tingkat menebak (TM), (7) standar kompetensi (SK), (8) kompetensi dasar (KD), dan indikator. Untuk menjaga dari hal-hal yang tidak diinginkan, nama siswa diwakilkan dengan atribut "Siswa kc-i". Rangkuman hasil kemampuan (θ) dan hasil nilai ulangan murni (NUM) pada uji coba disajikan dalam tabel berikut.

Berdasar hasil analisis data dari ke-51 siswa menunjukkan bahwa sistem inferensi yang dibangun dengan menggunakan metode pohon segitiga keputusan pada program CAT telah berhasil dengan baik, tepat, akurat, dan sesuai dengan kemampuan peserta dalam memilih butir-butir tes yang yang tepat dan sesuai dengan kemampuan peserta. Dalam hal itu, peserta dengan kemampuan tinggi menerima butir-butir tes dengan tingkat kesulitan tinggi. Peserta dengan kemampuan sedang menerima butir-butir tes dengan tingkat kesulitan sedang. Peserta dengan kemampuan rendah menerima butir-butir tes dengan tingkat kesulitan rendah. Disisi lain banyak butir yang diterima peserta antara yang satu dengan yang lain beragam (tidak sama). Peserta dengan kemampuan tinggi atau Peserta

dengan kemampuan rendah menerima banyak butir lebih sedikit dibandingkan dengan peserta berkemampuan sedang. Hal ini menunjukkan bahwa CAT yang dibangun dengan menggunakan pohon segitiga keputusan juga telah berhasil dengan baik dengan kemampuan peserta (*adaptif*).

Simpulan

Berdasarkan deskripsi data hasil penelitian dan pembahasan dapat diambil kesimpulan sebagai berikut: (1) Pengembangan CAT menggunakan metode pohon segitiga keputusan dalam prosedur pemilihan item selama pelaksanaan tes berdasarkan pada fungsi *fitness* terhadap kemampuan. *Software* CAT yang dikembangkan menggunakan bahasa pemrograman PHP (*Hypertext Preprocessor*) dan sistem basis data menggunakan My SQL (*Structured Query Language*). (2) Sistem inferensi model CAT menggunakan metode pohon segitiga keputusan mampu memilih butir-butir tes dengan tingkat kesulitan yang sesuai dengan benar-salah hasil respons peserta tes dan mampu mengatur banyak butir tes yang diberikan kepada peserta tes sesuai dengan tingkat kemampuannya. (3) Ketepatan dan keakuratan sistem inferensi model CAT dalam memilih butir-butir tes yang sesuai dengan kemampuan peserta tes diwujudkan dari hasil uji coba *beta* (siswa) dengan hasil bahwa siswa yang memiliki nilai ulangan murni (NUM) matematika tinggi di sekolah juga memiliki kemampuan (θ) tinggi, siswa yang memiliki nilai ulangan murni (NUM) matematika sedang di sekolah juga memiliki kemampuan (θ) sedang, dan siswa yang memiliki nilai ulangan murni (NUM) matematika rendah di sekolah juga memiliki kemampuan (θ) rendah. Nilai korelasi (r) antara hasil estimasi kemampuan (θ) pada CAT dengan nilai ulangan murni (NUM) matematika di sekolah siswa cukup tinggi yakni $r = 0,67$.

Program CAT yang dikembangkan ini berbasis internet maka pemanfaatannya bisa digunakan tes secara *online* sehingga bisa dimanfaatkan untuk ujian sekolah bahkan ujian nasional (UN) sebagai salah satu alternatif untuk meminimasi kecurangan dalam pelaksanaannya.

Daftar Pustaka

- Borg, W.R. & Gall, M.D. 2003. *Education research an introduction*. New York: Von Hoffman Press, Inc
- Djemari Mardapi. 2008. *Teknik penyusunan instrumen tes dan nontes*. Yogyakarta: Mitra Cendikia Press.
- Gronlund, N.E. 1976. *Measurement and evaluation in teaching*. New York : Macmillan Publishing Co
- Hambleton, R.K. & Swaminathan, H. & Rogers, H.J. 1991. *Fundamental of item response theory*. Newbury Park, CA: Sage Publication Inc.
- Lord, F.M. 1980. *Applications of item response theory to practical testing problems*, Hillsdale, NJ : Erlbaum.
- Masters, N.G. & Keeves, P.J. 1999. *Advances in measurement in educational research and assesment*. New York: Pergamon Press.
- Phankokkruad, M. & Woraratpanya, K. 2008. An automated decision system for computer adaptive testing using genetic algorithms. *Ninth ACIS International Conference on Engineering, Artificial Intellegence, Networking, and Parallel*, 655-660
- Wainer, H. 1990. *Computerized adaptive testing : A primer*. Hillsdale, NJ: Lawrence Erlbaum Associates, Publisher.
- Weiss, D.J. 2004. Computerized adaptive testing for effective and efficient measurement in counseling and education. *Measurement and Evaluation in Counseling and Development*, 37, 70.