

IMPLEMENTASI SISTEM PENILAIAN HASIL BELAJAR SISWA MENGGUNAKAN

COMPUTER ADAPTIVE TEST

Skripsi

diajukan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana Pendidikan Program Studi Pendidikan Teknik Informatika dan Komputer

Oleh



JURUSAN TEKNIK ELEKTRO FAKULTAS TEKNIK UNIVERSITAS NEGERI SEMARANG

2017

PERSETUJUAN PEMBIMBING

Nama : Khoirotun Nashihah

NIM : 5302412100

Program Studi : S1-Pendidikan Teknik Informatika dan Komputer

Judul Skripsi : Implementasi Sistem Penilaian Hasil Belajar Siswa

menggunakan Computer Adaptive Test (CAT)

Skripsi ini telah disetujui oleh pembimbing untuk diajukan ke sidang panitia ujian skripsi. Program Studi S1-Pendidikan Teknik Informatika dan Komputer FT UNNES.

Semarang, Desember 2016

Dosen Pembimbing I

Dosen Pembimbing II

Dra Dwi Purwanti, AhT, M.S. Drs

Drs. Djoko Adi Widodo, M.T.

NIP. 195910201990022001

NIP. 195909271986011001

PENGESAHAN

Skripsi dengan judul "Implementasi Sistem Penilaian Hasil Belajar Siswa Menggunakan
Computer Adaptive Test (CAT)" oleh Khoirotun Nashihah 5302412100, telah dipertahankan
di hadapan sidang Panitia Ujian Skripsi Fakultas Teknik Universitas Negeri Semarang pada
Rabu, 26 Oktober 2016.

Oleh

Nama

: Khoirotun Nashihah

NIM

: 5302412100

Program Studi

: S-1 Pendidikan Teknik Informatika dan Komputer

Panitia:

Ketua Panitia,

Sekreataris,

Dr.-Ing. Dhidik Prastiyanto, S.T., M.T.

NIP. 197805312005011002

Ir. Ulfah Mediaty Arief, M.T.

NIP. 196605051998622001

Penguji I

Penguji II

Penguji III

Drs. Slamet Seno Adi, M.Pd., M.T

Dra Dwi Purwanti, AhT, M.S.

Drs.Djoko Adi Widodo, M.T

NIP. 195812181985031004

NIP. 195910201990022001

NIP. 195909271986011001

Mengetahui

Dekan Fakultas Teknik UNNES

PERNYATAAN KEASLIAN

Saya menyatakan bahwa yang tertulis di dalam skripsi ini benar-benar karya sendiri, bebas dari plagiat, bukan jiplakan dari karya tulis orang lain, baik sebagian maupun seluruhnya. Pendapat atau temuan orang lain yang terdapat dalam skripsi dirujuk berdasarkan kode etik ilmiah. Apabila dikemudian hari terdapat plagiat dalam karya ini, maka saya bersedia menerima sanksi sesuai peraturan perundang-undangan.

Semarang, Desember 2016

Khoirotun Nashihah

NIM. 5302412109

UNIVERSITAS NEGERI SEMARANGI

MOTTO DAN PERSEMBAHAN

MOTTO

- Jika tidak bias membantu maka jangan menyusahkan selalu bermanfaat bagi orang lain ©
- Sesungguhnya Allah tidak akan merubah suatu kaum kalau kaum tersebut tidak punya kemauan dan usaha untuk merubah keadaan pada diri mereka sendiri. (Q.S Ar-Ra'du ayat 11).
- 3. Jadikanlah titik kelemahanmu menjadi titik kelebihanmu, teruslah mencoba dan berusaha, karena tidak ada usaha yang sia-sia.

PERSEMBAHAN

Skripsi ini saya persembahkan untuk:

- Bapak Suwanto dan Ibu Sukiyati yang selalu memberikan doa, dukungan, motivasi, semangat dan kasih sayang yang tiada henti selama ini.
- 2. Kakakku Nurul Azizah dan adik Ahmad Musthofa serta keluarga besarku terimakasih untuk dukungannya
- Sahabat-sahabatku seperjuangan dan teman-teman PTIK 2012 yang selalu membantu, terimakasih atas kekompakan dan waktu kebersamaannya.

ABSTRAK

Khoirotun Nashihah, 2016. Implementasi Sistem Penilaian Hasil Belajar Siswa Menggunakan *Computer Adaptive Test* (CAT). *Skripsi*. Program Studi Pendidikan Teknik Informatika dan Komputer Jurusan Teknik Elektro Fakultas Teknik Universitas Negeri Semarang. Pembimbing: Dra Dwi Purwanti, AhT, M.S. dan Drs. Djoko Adi Widodo, M.T.

Perkembangan teknologi informasi sekarang ini sudah sangat pesat dengan adanya berbagai macam alat bantu yang memudahkan manusia. Dalam bidang pendidikan salah *satunya* adalah bidang evaluasi. Salah satu untuk mengembangan metode tes ini adalah dengan menggunakan CAT (*computer adaptive test*). adaptif disini memiliki pengertian bahwa butir soal (tes) yang diberikan disesuaikan dengan tingkat kemampuan setiap peserta tes. Tujuan dari penelitian ini adalah penerapkan sistem penilaian dengan menggunakan *computer adaptive test* dan mengetahui sistem yang dikembangkan efektif jika digunakan sebagai sistem penilaian hasil belajar.

Pengembangan aplikasi *computer adaptive test* ini menggunakan metode *waterfall*. Pengujian sistem dilakukan dengan dua tahap yaitu pengujian alpha dan pengujian beta. Pengujian alpha merupakan pengujian fungsional yang digunakan untuk menguji sistem yang telah dibuat dengan metode pengujian *black box* dan *white box*. Pengujian beta difokuskan pada aspek/variabel penelitian kepraktisan dan keefektifan sistem jika akan diterapkan. Pengambilan data dengan kuesioner. Analisis data menggunakan *skala Likert* untuk menentukan panjang interval penilaian untuk sistem.

Dari hasil pengujian penilaian hasil belajar menggunakan *Computer Adaptive Test* (CAT) berada pada kategori baik dengan nilai rata-rata 81,25%. sistem panilaian menggunakan CAT efektif untuk digunakan. Namun keterbatasan penelitian tersebut hanya dalam skala yang terbatas di mata pelajaran MySQL di SMK Negeri 10 Semarang.

Kata Kunci: Sistem Penilaian, Computer Adaptive Test, Efektivitas, Waterfall.

LIND/ERSITAS NEGERI SEMARANG.

KATA PENGANTAR

Puji syukur Alhamdulillah penulis panjatkan kehadirat Allah SWT yang telah melimpahkan nikmat, rahmat, taufik dan hidayah-Nya, sehingga penulisan skripsi ini dapat diselesaikan dengan baik.

Laporan skripsi ini disusun guna memenuhi salah satu syarat untuk menyelesaikan Program Studi Pendidikan Teknik Informatika dan Komputer Jurusan Teknik Elektro Fakultas Teknik Universitas Negeri Semarang.

Dalam menyusun skripsi ini tidak terlepas dari bantuan, dorongan, serta doa dari berbagai pihak baik secara langsung maupun tidak langsung. Pada kesempatan ini, dengan segala kerendahan hati penulis menyampaikan terima kasih yang setulus-tulusnya kepada:

- 1. Dra Dwi Purwanti, AhT, M.S. Dosen Pembimbing I dan Drs. Djoko Adi Widodo, M.T. Dosen Pembimbing II yang telah sabar memberi bimbingan.
- 2. Keluarga tercinta, Bapak, Ibu, Kakak dan Adik, Sahabatku Anita, Rendi, Anggi, Annas, Safira, Diah, Alfini, Fatwa yang selalu memberikan semangat dan membantu selama *study* kuliah ini, serta teman-teman PTIK 2012 yang tidak dapat saya sebut satu per satu.
- 3. Semua pihak yang membantu terselesaikannya skripsi

Semarang, Desember 2016

Penulis

DAFTAR ISI

| | | Halaman |
|--------|--|---------|
| HALAN | AAN JUDUL | i |
| LEMBA | AR PERSETUJUAN PEMBIMBING | ii |
| PENGE | SAHAN | iii |
| LEMBA | AR KEASLIAN | iv |
| MOTTC | D DAN PERSEMB <mark>A</mark> HAN | v |
| ABSTR. | AK | vi |
| KATA I | PENGA <mark>NTAR</mark> | vii |
| DAFTA | R ISI | viii |
| DAFTA | R TA <mark>BEL</mark> | xi |
| DAFTA | R GAMBAR | xii |
| | R LAMPIRAN | |
| BAB I | PENDAHULUAN | 1 |
| 1.1 | Latar Belakang | 1 |
| 1.2 | Identifikasi Masalah | 3 |
| 1.3 | Identifikasi MasalahRumusan Masalah | 4 |
| 1.4 | Batasan Masalah S. L.A.S. N.L. C. B. B. C. L. MAGARAGE | |
| 1.5 | Penegasan Istilah | 4 |
| 1.6 | Tujuan Penelitian | 5 |
| 1.7 | Manfaat Penelitian | 5 |
| BAB II | KAJIAN PUSTAKA | 7 |
| 2.1 | Paper Based Test (PBT) | 7 |
| 2.2 | Computer Based Test (CBT) | 7 |
| 2.3 | Computer Adaptive Test (CAT) | 8 |

| | | На | alaman |
|---------|-------|---|--------|
| 2 | 2.3.1 | Teori Tes Klasik | 9 |
| 2 | 2.3.2 | Item Response Theory | 12 |
| 2 | 2.3.3 | Konsep Computer Adaptive Test (CAT) | 15 |
| 2.4 | Renc | cana Pengembangan CAT | 19 |
| 2.5 | Нуре | ertext Prepocessor(PHP) | 20 |
| 2.6 | MyS | QL | 21 |
| 2.7 | Pene | litian Yang Relevan | 23 |
| 2.8 | Kera | ngka Berfiki <mark>r</mark> | 25 |
| BAB III | MET | TODE <mark>PENELI</mark> TIAN | 26 |
| 3.1 | Wak | tu <mark>d</mark> an <mark>Temp</mark> at Pelaksanaan | 26 |
| 3.2 | Desa | in Penelitian | 26 |
| 3.3 | Alat | dan Bahan Penelitian | 54 |
| 3 | 3.3.1 | Alat Penelitian | 54 |
| 3 | 3.3.2 | Bahan Penelitian | 54 |
| 3.4 | Parai | meter Pene <mark>litian</mark> | 55 |
| 3.5 | Tekn | nik Pengum <mark>pulan</mark> Data | 56 |
| 3.6 | Kalib | brasi Instrum <mark>en</mark> | 56 |
| 3.7 | Tekn | nik Analisa Data | 58 |
| BAB IV | | SIL DAN PEMBAHASAN | |
| 4.1 | Desk | cripsi Data | 61 |
| 4 | 1.1.1 | Hasil Sistem Penilaian dengan Computer Adaptive Test | 61 |
| 4 | 1.1.2 | Hasil Pengujian Respon Pengguna | 62 |
| 4.2 | Peml | bahasan | 72 |
| BAB V | PENU | UTUP | 75 |
| 4.3 | Kesi | mpulan | 75 |
| 4.4 | Sarai | n | 75 |
| DAFTA | R PUS | STAKA | 77 |
| LAMPII | RAN | | |

DAFTAR TABEL

| | Halaman |
|--|---------|
| Tabel 2.1 Tabulasi perhitungan Computer Adaptive Test | 17 |
| Tabel 3.1 Pengujian Sistem | 48 |
| Tabel 3.2 pengujian <i>black-box</i> | 51 |
| Tabel 3.3 Kisi-kisi Instrumen | 57 |
| Tabel 3.4 Interval pengkategorian skor kriteria kualitatif | 60 |
| Tabel 4.1 Respon pengguna | 63 |
| Tabel 4.2 hasil aspek nilai sistem | 65 |
| Tabel. 4.3 Hasil penilaian aspek tampilan sistem | 67 |
| Tabel. 4.4 Has <mark>il penilaian aspe</mark> k <mark>navi</mark> gasi | 69 |
| Tabel 4.5 Hasil penilaian keandalan aplikasi | 71 |



DAFTAR GAMBAR

| | | Halaman |
|-------------|--|---------|
| Gambar 2.1 | Kurva model 1 parameter | 14 |
| Gambar 3.1 | Model Sekuensial Linier | 27 |
| Gambar 3.2 | Desain Penelitian | 27 |
| Gambar 3.3 | Diagram konteks | 29 |
| Gambar 3.4 | DFD level 1 | 30 |
| Gambar 3.5 | DFD level 2 proses 1 | 31 |
| Gambar 3.6 | DFD level 2 proses 2 | 31 |
| Gambar 3.7 | DFD level 2 proses 3 | 32 |
| Gambar 3.8 | DFD level 2 proses 3 | 32 |
| Gambar 3.9 | Activity diagram admin | 33 |
| Gambar 3.10 | Activity diagram peserta | 34 |
| Gambar 3.11 | Rancangan Database | 35 |
| Gambar 3.12 | Diagram Perancangan Menu Peserta/Siswa | 36 |
| Gambar 3.13 | Diagram Perancangan Menu Admin/guru | 36 |
| Gambar 3.14 | Halaman login peserta | 37 |
| Gambar 3.15 | Halaman utama siswa | 37 |
| Gambar 3.16 | Halaman memilih tes | 38 |
| Gambar 3.17 | Halaman tampilan soal | 38 |
| Gambar 3.18 | Tampilan hasil tes siswa | 39 |
| Gambar 3.19 | Tampilan login admin | 39 |
| Gambar 3.20 | halaman menu soal | 40 |

| Gambar 3.21 Tampilan form tambah soal | 40 |
|--|-----|
| Gambar 3.22 Tampilan halaman menu peserta | 41 |
| Gambar 3.23 Tampilan form tambah peserta | 41 |
| Gambar 3.24 Tampilan halaman menu tes | 42 |
| Gambar 3.25 Tampilan form tambah tes | 42 |
| Gambar 3.26 Tampilan halaman hasil tes | 43 |
| Gambar 3.27 Halaman has <mark>il</mark> tes me <mark>nampi</mark> lkan sis <mark>wa ya</mark> ng mengikuti tes | 43 |
| Gambar 3.28 Halaman hasil tes menampilkan detail tes dari siswa yang dipilih. | 44 |
| Gambar 3.29 <i>U<mark>se case diagram</mark></i> | 45 |
| Gambar 3.30 <mark>Pengubahan <i>Flowchart</i> menjadi <i>Flow Graph</i> Algoritma <i>Compu</i></mark> | ter |
| Adaptive test | 49 |
| Gambar 3.31 Path White-box Testing | 50 |
| Gambar 4.1 Hasil tes m <mark>enggu</mark> nakan <i>Compute<mark>r Ada</mark>ptive Test</i> | 62 |
| Gambar 4.2 Grafik Respon Pengguna Guru | 64 |
| Gambar. 4.3 grafik penilaian aspek isi sistem | 66 |
| Gambar 4.4 Grafik penilaian aspek tampilan sistem | 68 |
| Gambar. 4.5 grafik penilaian aspek navigasi sistem | 70 |
| Gambar 4 6 grafik penilaian keandalan sistem | 71 |

DAFTAR LAMPIRAN

| | | Halaman |
|-------------|--|---------|
| Lampiran 1 | Formulir Usulan Topik Skripsi | 80 |
| Lampiran 2 | Surat Usulan Pembimbing | 81 |
| Lampiran 3 | Surat Keputusan Penetapan Dosen Pembimbing | 82 |
| Lampiran 4 | Surat Ijin Penelitian Dinas Pendidikan Kota Semarang | 83 |
| Lampiran 5 | Surat Selesai Penelitian | 84 |
| Lampiran 6 | Soal Uji Coba | 85 |
| Lampiran 7 | Hasil Uji coba soal | 93 |
| Lampiran 8 | Tampilan Sistem | 94 |
| Lampiran 9 | Angket Respon Pengguna(siswa) | 103 |
| Lampiran 10 | Rekap hasil angket Respon Pengguna(gutu) | 104 |
| Lampiran 11 | Dokumentasi | 105 |



BABI

PENDAHULUAN

1.1 Latar belakang

Tes adalah salah satu alat yang paling banyak digunakan dalam pendidikan untuk mengukur tingkat kemampuan siswa. Kemajuan teknologi komputer dan teori pengujian telah mempercepat perubahan format ujian dari konvensional ke pengujian berbasis komputer. Sekarang ini dikenal tes dengan CBT (computer based test). CBT pada prinsipnya seperti ujian menggunakan kertas dan pensil biasa, hanya saja penyajiannya menggunakan komputer. Semua peserta tes dalam CBT mengerjakan soal yang sama. Untuk meningkatkan efisiensi dan akurasi dalam mengukur kemampuan peserta tes, dikembangkan ujian CBT yang bersifat adaptif yaitu pengujian berbasis komputer lainnya adalah Computer Adaptive Test (CAT). menurut Lord dalam penelitian Agus Santoso(2010), adaptif disini memiliki pengertian bahwa butir soal (tes) yang diberikan disesuaikan dengan tingkat kemampuan setiap peserta tes.

CAT didasarkan pada item response theory(IRT). Menurut Hambleton dalam penelitian Rukli(2011) pada IRT, karakteristik soal yang mengacu pada parameter independen terhadap kelompok uji(sample free) sehingga cocok dengan tes adaptif. Ada tiga model logistik dalam teori respon butir, yaitu model logistik satu parameter (1PL), model logistik dua parameter (2PL), dan model logistik tiga parameter(3PL). perbedaan dari model tersebut terletak pada banyaknya parameter yang digunakan.

Parameter tersebut adalah indeks kesukaran, indeks daya beda butir dan indeks tebakan semu.

Perkembangan CAT di Indonesia masih dalam tahap penelitian dan uji coba. Misalnya penelitian yang dilakukan oleh Haryanto(2009) dengan menggunakan algoritma logika fuzzy dalam memilih soal-soal yang sesuai dengan kemaampuan peserta te dengan menggunakan model logistic dua parameter (2PL). Penelitian Santoso(2010) menggunakan model logistic 3PL pada ujian masuk Universitas Terbuka. Penelitian Rukli (2011) mengunakan model logistic 3PL yang diterapkan sebagai sistem pendukung keputusan dalam pengujian Computerized Adaptive Testing. Selain itu beberapa negara juga telah menggunakan CAT dalam sistem pengujian. Misalnya, Jepang telah meluncurkan CAT secara online dengan nama J-CAT pada http://www.j-cat.org/ disebutkan Imai dalam penelitian Rukli(2011). Hasil penelitian tersebut menunjukan bahwa CAT dapat dijadikan media pengujian yang baik karena akurasi yang tinggi, baik dari jumlah soal yang sedikit, kerjasama antar peserta tes yang dapat dikurangi, dan peserta tes dapat mengetahui kemampuan secara langsung.

Berdasarkan penelitian Erawan(2009), ada beberapa keuntungan penggunaan CAT dalam sistem ujian. Pertama adalah waktu yang diperlukan untuk ujian lebih singkat, peserta didik dalam mengerjakan soal bersifat individual. Hal ini berarti bahwa butir yang disajikan untuk tiap peserta didik berbeda, sehingga mengurangi peluang bekerjasama. Namun hasil tes bisa dibandingkan karena semua butir soal dalam bank soal telah dikalibrasi.

Melihat kondisi tersebut dalam penelitian ini bermaksud mengembangkan inovasi pada ujian berbasis komputer di indonesia, CBT ke ujian adaptif berbasis komputer, *computer adaptive test*. Dengan menerapkan CAT sebagai sistem penilaian hasil belajar siswa.

Oleh karena itu, skripsi ini diberi judul "IMPLEMENTASI SISTEM PENILAIAN HASIL BELAJAR SISWA DENGAN MENGGUNAKAN COMPUTER ADAPTIVE TEST"

1.2 Identifikasi Masalah

Setelah memperhatikan kondisi tersebut dapat diidentifikasi beberapa masalah antara lain:

- 1. Pada *computer based test* (CBT) pengukuran kemampuan peserta tes menggunakan *True Score*. Sementara kemampuan peserta tes sebenarnya bergantung pada isi tes.
- 2. Perkembangan *Computer Adaptive Test* (CAT) di indonesia terbilang masih tertinggal dibandingkan dengan negara lain seperti Amerika yang menggunakan CAT untuk tes kepegawaian dan mengukur kemampuan mahasiswa di perguruan tinggi, masih sedikit lembaga yang mengembangkan CAT di Indonesia.

1.3 Rumusan Masalah

Berdasarkan latar belakang di atas maka rumusan masalah pada penelitian ini adalah bagaimana mengembangkan sistem penilaian hasil belajar menggunakan konsep *Computer Adaptive test*?

1.4 Batasan Masalah

Pada skripsi ini, masalah yang di teliti akan dibatasi pada:

- 1. Sistem ya<mark>ng dikemb</mark>angkan hanya merupakan alat bantu evaluasi belajar.
- 2. Sistem yang dikembangkan adalah berbasis web dengan dukungan PHP programming dan database MySQL.

1.5 Penegasan Istilah

Istilah yang digunakan dalam skripsi ini adalah sebagai berikut:

- 1. Implementasi : penerapan suatu hal pada aksi, aktivitas tindakan serta adanya mekanisme untuk mencapai tujuan tertentu. . (Usman:2002)
- 2. Computer Adaptif Test (CAT): suatu konsep tes berbasis komputer dimana butir soal (tes) yang diberikan disesuaikan dengan tingkat kemampuan setiap peserta tes. Menurut Lord dalam penelitian Agus Santoso(2010).
- 3. Sistem penilaian: suatu cara yang digunakan untuk melaksanakan penilaian mencakup kegiatan mengumpulkan menganalisa serta mengintreprestasikan informasi yang dapat digunakan untuk membuat kasimpulan tentang karakteristik seorang atau objek. Gronlund & Linn dalam Kusaeri (2012:8)

4. Hasil Belajar: Perolehan dari proses belajar. Hasil belajar dapat berupa perubahan kemampuan kognitif,afektif,dan psikomotorik yang mempengaruhi perubahan perilaku peserta didik.

Sehingga dapat disimpulkan untuk "Implementasi Sistem Penilaian Hasil Belajar dengan Menggunakan *Computer Adaptive Test*" adalah Suatu tindakan untuk melaksanakan penilaian dari proses belajar dengan menggunakan suatu konsep tes berbasis komputer dimana butir soal (tes) yang diberikan disesuaikan dengan tingkat kemampuan setiap peserta tes.

1.6 Tujuan Penulisan

Berdasarkan uraian rumusan masalah diatas, tujuan dilakukannya penelitian ini adalah untuk mengembangkan sistem penilaian hasil belajar menggunakan konsep computer adaptive test.

1.7 Manfaat penulisan

Hasil penelitian ini diharapkan dapat bermanfaat bagi semua pihak, diantaranya sebagai berikut:

1. Bagi peneliti

penelitian ini memberikan manfaat kepada peneliti karena peneliti akan mengembangkan kemampuannya dalam merancang dan membangun sebuah aplikasi serta sebagai dasar pembelajaran untuk penelitian yang selanjutnya.

2. Bagi guru

Hasil dari penelitian ini dapat digunakan sebagai alat tes. Sehingga dapat digunakan sebagai alternative penilaian. Mendukung aktivitas pembelajaran dengan memanfaatkan teknologi informasi.

3. Bagi peserta didik

Hasil dari penelitian ini dapat memberikan pengalaman tentang ujian berbasis komputer.

4. Bagi lembaga pendidikan

Penelitian ini dapat memberikan dukungan dalam rangka meningkatkan mutu dan kualitas sistem penilaian.



BAB II

KAJIAN PUSTAKA DAN LANDASAN TEORI

2.1 Paper Based Test (PBT)

Tes adalah alat atau instrumen yang dapat digunakan dalam kegiatan evaluasi. Istilah tes tidak hanya popular di lingkungan persekolahan, tetapi juga diluar sekoalh bahkan di masyarakat umum. Kita sering mendengar istilah tes kesehatan, tes olahraga, tes makanan dan sebagainya. Berbagai jenis tes yang berkembang, salah satu tes yang paling umum berdasarkan media yang digunakan adalah *paper based test (PBT)*. PBT yang diselenggarakan secara tertulis dan menggunakan kertas. Semua soal disajikan di dalam kertas dan menjawab soal juga megguanakan kertas (LJK).

2.2 Computer Based Test (CAT)

Kemajuan teknologi dalam empat puluh tahun terakhir, yang ditandai dengan munculnya perangkat komputer telah merevolusi hampir setiap aspek dalam pengujian psikologis dan penilaian, (Drasgow & Mattern, 2006, p.59). Hal ini dapat dipahami karena perkembangan teknologi komputer pada generasi baru memungkinkan penyimpanan butir tes dalam sebuah data base bank soal dalam jumlah yang besar. Model Pengujian Berbasis Komputer (CBT) diselenggarakan dengan menggunakan komputer. Karakteristik dari tes ini sama dengan tes tertulis biasanya yaitu menggunakan satu perangakat tes untuk beberapa peserta dengan

panjang waktu tes yang sama. Perbedaannya terletak pada teknik penyampaian (*delivery*) butir soal yang tidak lagi meggunakan kertas (*paper*), baik untuk naskah soal maupun lembar jawaban soalnya. Sistem skoring atau koreksi langsung dilakukan oleh komputer.

Komputer mampu melakukan perhitungan yang rumit dengan cepat sehingga mampu menampilkan butir tes sesuai dengan jawaban peserta tes, bahkan mampu memberikan saran cerdas kepada pengguna berdasarkan basis pengetahuan yang ada pada program yang dijalankan dalam komputer tersebut.

2.3 Computer Adaptive Test (CAT)

Berbagai usaha telah dilakukan para ahli pengukuran pendidikan untuk mengatasi masalah dalam penilaian hasil belajar. Prosedur penyusunan soal kriteria soal yang baik yang baik telah dikembangkan. Rumus dan program komputer untuk menentukan parameter soal atau butir soal yang telah dibuat. Untuk memperbaiki proses pelaksanaan penilaian hasil belajar, para ahli pengukuran pendidikan bekerja sama dengan programer komputer telah memanfaatkan komputer untuk melaksanakan penilaian hasil belajar. Dengan memanfaatkan komputer dalam penilaian hasil belajar, maka kerjasama dan kecurangan yang dilakukan peserta tes dapat dikurangi; selain itu pemberian skor atau nilai kepada peserta tes akan objektif dan konsisten.

CAT memberikan sejumlah butir tes dengan tingkat kesulitan yang disesuaikan dengan kemampuan peserta tes. Dalam penerapan analisis data hasil uji coba soal

untuk mengetahui secara empiris. Ada dua pendekatan teori tes klasik dan menggunakan teori respon butir:

2.3.1 Teori Tes Klasik

Teori tes klasik, atau disebut juga teori tes skor murni klasik didasarkan pada model aditif, yaitu skor amatan merupakan penjumlahan dari skor yang sebenarnya dan skor kesalahan pengukuran menurut Allen Yen dalam Hadi(2013:2). Asumsi yang mendasari teori tes klasik ini dijadikan dasar mengembangkan rumus-rumus matematis untuk mengestimasi validitas dan koefisien reliabilitas pada perangkat tes yang digunakan untuk menilai kualitas tes. Kualitas tes dalam teori tes klasik juga dapat ditentukan dengan indeks kesukaran dan daya pembeda.

a. Tingkat Kesukaran

Tingkat kesukaran yang disimbolkan dengan *p* merupakan salah satu parameter butir soal yang sangat berguna dalam analisis soal. Tingkat kesukaran dapat dihitung dengan berbagai cara, yaitu (a) skala kesukaran linier, (b) skala kesukaran bivariate, (c) indeks Davis, dan (d) proporsi menjawab benar (Bahrul Hayat, 1999). Secara metematis tingkat kesukaran yang dihitung dengan proporsi menjawab benar dirumuskan dengan:

$$P = \frac{\sum B}{N}$$

(Samsul Hadi, 2013:3)

Keterangan =

B : Banyaknya peserta yang menjawab benar

N : Jumlah peserta yang menjawab

Setelah menghitung indeks kesukaran untuk setiap butir soal, soal-soal tersebut dapat digolongkan ke dalam soal kategori mudah, sedang, dan sukar. Penggolongan soal-soal tersebut berdasarkan kriteria indeks kesukaran soal. Kriteria indeks kesukaran soal menurut Arikunto (2009:208) yaitu:

P : 0.00 - 0.30 = soal kategori sukar

P : 0.30 - 0.70 = soal kategori sedang

P : 0.70 - 1.00 = soal kategori mudah

b. Daya Beda Soal

Daya pembeda adalah kemampuan sesuatu soal untuk membedakan antara siswa yang pandai (berkemampuan tinggi) dengan siswa yang bodoh (berkemampuan rendah) (Arikunto,2009:211). Tes dikatakan memiliki daya pembeda jika tes diujikan kepada anak yang berprestasi tinggi, maka hasilnya tinggi. Jika diujkan kepada siswa berprestasi rendah, maka hasilnya rendah (Sudjana,2009:141). Angka yang menunjukkan besarnya daya pembeda disebut indeks diskriminasi. Indeks diskriminasi berkisar 0,00 sampai 1,00 dan mengenal tanda positif dan negatif (Arikunto,2009:211). Rumus untuk menentukan indeks diskriminasi adalah:

$$DP = \frac{BA}{IA} - \frac{BS}{IS}$$

Keterangan:

J : jumlah peserta tes

J_A : banyaknya peserta kelompok atas

J_B : banyaknya peserta kelompok bawah

 $B_{\mbox{\scriptsize A}}$: banyaknya peserta kelompok atas yang menjawab soal dengan benar

 B_{B} : banyaknya peserta kelompok bawah yang menjawab soal itu dengan benar

(Arikunto, 2009:214)

Berdasarkan hasil perhitungan indeks diskriminasi setiap butir soal, dapat diketahui klasifikasi daya pembeda dari soal tes yang diuji cobakan. Soal mempunyai klasifikasi daya pembeda jelek, cukup, baik, baik sekali, dan tidak baik.

Klasifikasi daya pembeda butir soal antara lain, yaitu:

D : 0.00 - 0.20 = jelek(poor)

D : 0.20 - 0.40 = cukup (satisfactory)

D : 0.40 - 0.70 = baik (good)

D : 0.70 - 1.00 = baik sekali (excellent)

D : negatif, semuanya tidak baik. Jadi, semua butir soal yang mempunyai nilai D negatif sebaiknya dibuang saja

(Arikunto,2009:218)

2.3.2 Item Response Theory (IRT)

Menurut Subagyo(2014) dalam penelitiannya bahwa "IRT (*item response theory*) merupakan model matematika yang memperhitungkan kemungkinan orang memberikan jawaban dengan benar untuk setiap item."

Ada tiga parameter logistic dalam teori respons butir, yaitu model logistik satu parameter (1PL), model logistic dua parameter (2PL) dan model logistik tiga parameter (3PL). perbedaan 3 model tersebut terletak dari banyaknya parameter yang digunakan untuk menggambarkan karakteristik butir dalam model yang digunakan. Parameter-parameter tersebut adalah indeks kesukaran, indeks daya beda butir dan indeks tebakan semu. Secara matematis, model logistik tiga parameter (3PL) dapat dinyatakan sebagai berikut.

$$P_i(\theta) = \frac{(1 - C_i)e^{Dai(\theta - b_i)}}{1 + e^{Dai(\theta - b_i)}}$$

Keterangan:

θ : Tingkat Kemampuan (ability) peserta tes

 $Pi(\theta)$: Probabilitas peserta tes yang memiliki kemampuan θ dapat menjawab butir i dengan benar

ai :indeks daya pembeda

bi :indeks kesukaran butir ke i

ci :indeks tebakan semu butir ke-i

e : bilangan natural yang nilainya mendekati 2,718

D: Faktor penskala yang harganya 1,7

Kurva karakteristik butir soal yang dianalisis dengan model 3 parameter logistik yang memiliki a=1,117; b=2,287; dan c=0,198 ditunjukan pada Gambar 1.1. Gambar ini menunjukan bahwa probalitas menjawab benar tidak berawal dari 0, tetapi berawal dari 0,198. Jadi jawaban yang bersifat tebakan mempunyai kemungkinan benar 19,8%.

Model yang kedua adalah model 2 parameter merupakan bagian dari model 3 parameter, dan merupakan kasus khusus dari model 3 parameter, yakni ketika c=0. Secara matematika dapat dirumuskan sebagai berikut:

$$P_i(\theta) = \frac{e^{Dai(\theta - bi)}}{1 + e^{Dai(\theta - bi)}}$$

Seorang siswa dalam menjawab soal akan memiliki probabilitas menjawab benar dengan nilai tertentu. Untuk Suatu tes dengan soal yang hanya memiliki satu parameter, parameter kesulitan, nilai probabilitas ini dapat dihitung dengan

$$Pi(\theta) = \frac{e^{(\theta - bi)}}{1 + e^{(\theta - bi)}}$$

Keterangan:

 $\operatorname{Pi}(\Theta)$: probabilitas seorang siswa menjawab benar suatu soal dengan paramater b

Θ : kemampuan terestimasi dari siswa

b : parameter kesulitan yang menentukan tingkat kesulitan soal

Rumusan ini ketika proses ujian berlangsung merupakan suatu proses iterasi. Untuk setiap soal yang diberikan, akan didapatkan suatu nilai probabilitas tersendiri. Dari rumusan ini dapat digambarkan suatu grafik yang menghubungkan antara probabilitas dan kemampuan siswa.



Gambar 2.1 kurva model 1 parameter

Kurva karakteristik butir soal yang dianalisis dengan model 1 parameter logistik ditujukan pada gambar diatas. Besarnya a=1 (konstanta), sedangkan besarnya b=0,018.

Di dalam menerapkan *item response theory* (IRT) ke permasalahan pengukuran, asumsi umum bahwa kemampuan peserta tes tercermin dalam pencapaian dalam menyelesaikan item. Asumsi ini terjadi pada hampir semua aplikasi pengujian adaptif. Model IRT paling sesuai untuk pengetesan adaptif adalah fungsi tiga

parameter logistik (3-PL). Alasan yang utama untuk memilih model fungsi tiga parameter logistik (3-PL) adalah sesuai dengan data item pilihan ganda yang lebih baik daripada model fungsi satu parameter logistik (1-PL) atau model fungsi dua parameter logistik (2-PL).

Fungsi informasi item sangat berperan di dalam pengetesan adaptip. Item yang memberikan secara maksimal kepada ketepatan pengukuran, terpilih untuk ditampilkan di dalam monitor. Item yang menyediakan informasi banyak adalah, item yang berada di atas kemampuan penempuh tes (kira-kira) 50%-60% kesempatan menjawab dengan tepat.

Dalam penelitian Suwarno(2009), model item respon sangat sesuai untuk pengujian yang adaptif sebab kemungkinan untuk memperoleh estimasi kemampuan peserta tes tidak terikat pada satuan item tes tertentu. Pengujian yang adaptip akan selalu mengarah dalam suatu wilayah *item response theory* (IRT). Walaupun masingmasing peserta tes menerima item yang berbeda-beda, kesukaran item yang berbeda, namun IRT memungkinkan untuk membandingkan estimasi kemampuan dari peserta tes yang berbeda-beda.

UNIVERSITAS NEGERI SEMARANG

2.3.3 Konsep Computer Adaptive Test (CAT)

CAT hanya dapat dilakukan dengan komputer. Komputer dapat menyimpan informasi tes (item-item tes dan karakteristiknya = item bank) yang cukup besar dan dapat menampilkan item-item yang sesuai dengan kemampuan peserta tes. Komputer juga dapat mengatur, dan memberikan skor tes pada setiap peserta tes sesuai dengan

potensi masing-masing peserta tes dengan cepat dan akurat menurut penelitian Suwarto (2009) Di dalam *computerized adaptive testing* (CAT), urutan item yang diatur untuk dimunculkan ke monitor tergantung pada pencapaian peserta tes pada item sebelumnya. Atas dasar pencapaian kemampuan dari peserta tes tersebut, maka item yang mempunyai informasi maksimal tentang tingkatan kemampuan peserta tes tersebut diatur oleh komputer. Dengan demikian maka, panjang tes memungkinkan untuk diperpendek dengan tanpa menghilangkan ketepatan pengukuran.

Setelah peserta tes menyelesaikan beberapa item tes (dua atau tiga item) yang diperkenalkan pada awal tes, estimasi untuk kemampuan awal penempuh tes dapat diperoleh. Komputer diprogramkan untuk memilih item dari item bank yang tersedia dan memberikan informasi yang banyak tentang kemampuan peserta tes, semuanya ini dilakukan oleh komputer berdasarkan pada estimasi awal. Rincian bagaimana item tes dapat dipilih dan kemampuan peserta tes dapat diestimasi, dapat dijelaskan sebagai berikut: Pemberian item kepada peserta tes sampai sejumlah item yang telah diatur dan ditetapkan, seberapa tingkat ketepatan pengukuran yang diinginkan (yaitu, standard error = SE) tentang estimasi kemampuan yang dicapai.

Rumus yang digunakan untuk menghitung kemampuan (θ) probabilitas menjawab benar berdasarkan kemampuan tersebut $(P(\theta))$, probabilitas menjawab salah $(Q(\theta))$, fungsi informasi butir $I_i(\theta)$, dan kesalahan baku $SE(\theta)$ adalah

$$\theta = bi + \frac{1}{Da_i} \ln (0.5(1 + \sqrt{1 + 8c_i}))$$

$$P_i(\theta) = \frac{(1 - C_i)e^{Dai(\theta - b_i)}}{1 + e^{Dai(\theta - b_i)}}$$

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

$$I_i(\theta) = P_i(\theta)Q_i(\theta)$$

$$SE(\theta) = \frac{1}{\sqrt{\sum_{i=1}^{N} I_i(\theta)}}$$

Keterangan:

θ : Kemampuan

 $P(\theta)$: Probabilitas menjawab benar

 $Q(\theta)$: Probabilitas menjawab salah

 $I(\theta)$: Fungsi informasi butir

 $SE(\theta)$: Kesalahan Baku

Proses tersebut dilakukan berulangulang sampai stopping rule tercapai, kemudian kemampuan akhir (θ) peserta tes dihitung. Hasil perhitungan table tersebut dapat ditabulasi seperti tabel 2.1.

Tabel 2.1 Tabulasi perhitungan Computer Adaptive Test

| | | | | | | θ | θSetelah | $P(\theta)$ | $Qi(\theta)$ | IIF | $SE(\theta)$ | Selisih |
|----|-------|------|-----------|---------|---------|-------|----------|-------------|--------------|------|--------------|---------|
| | | | | 1113 | | Awal | jawab | 1(0) | Q1(0) | *** | SE(0) | SE |
| | Kode | Daya | Tingkat | | Skor | | _ | | | | | Antar |
| No | Butir | Beda | kesulitan | Tebakan | jawaban | | | | | | | Iterasi |
| 1 | 93 | 0.13 | 0.101.01 | 0.3 | 0 | 0.00 | 0.00 | 0.50 | 0.50 | 0.25 | 2.00 | 2.00 |
| 2 | 87 | 0.61 | -0.45 | 0.81 | 1 | 0.00 | -0.45 | 0.50 | 0.50 | 0.25 | 1.41 | 0.59 |
| 3 | 6 | 0.6 | -0.25 | 0.13 | 0 | -0.45 | -0.45 | 0.58 | 0.42 | 0.24 | 1.16 | 0.25 |

Pengembangan CAT memerlukan evaluasi pada enam komponen:

a. Model respon butir

Menggunakan respon butir soal model logistic dua parameter, yaitu tingkat kesulitan dan daya beda soal.

b. Bank soal

Soal yang sudah dianalisis dijadikan bank soal.

c. Pemilihan butir soal awal

Peserta diberikan soal dengan tingkat kesulitan awal yang sedang yaitu batas tengah 0-1 yaitu 0.5. Karena tingkat kesulitan awal (b awal) =0.5 dianggap mempunyai tingkat kemampuan awal (θ awal) sedang (θ awal=0).

d. Prosedur pemilihan butir soal

Setelah diberikan tingkat kesulitan sedang dapat dijawab benar, peserta diberikan soal baru dengan tingkat kesulitan yang lebih tinggi, sebaliknya jika dijawab salah maka peserta diberi soal yang lebih mudah kemudian kemampuan (θ) setelah menjawab soal baru $P(\theta)$, $Q(\theta)$, $I_i(\theta)$, $SE(\theta)$ dan harga mutlak selisih kesalahan baku antar penyajian soal dihitung. Proses ini diulang-ulang sampai stopping rule tercapai.

e. Aturan pemberhentian

Jika tingkat kestelitian estimasi kemampuan telah tercapai yang ditandai dengan harga mutlak selisih SE antar iterasi sangat kecil (<=0,01) maka penyajian

soal dihentikan. Dalam hal ini kemampuan (θ) peserta tes adalah kemampuan tertinggi yang pernah dicapai.

f. Penilaian kwantitatif

Supaya hasil CAT lebih mudah diterima oleh banyak pihak, kemampuan peserta tes tersebut diubah menjadi skor dengan skala terendah 0 dan tertinggi 100. Dengan menggunakan rumus sebagai berikut:

Skor(100)=
$$50+50*(\theta \text{ akhir})$$

Menurut Samsul Hadi (2013:12), Computerized adaptive testing (CAT) adalah sistem pengujian berbantuan computer yang lebih maju dibandingkan CBT. Dalam CAT butir soal yang diberikan kepada peserta tes disesuaikan dengan kemampuan peserta tes. Proses penyajian butir soal dalam CAT dilakukan secara berulang kali sampai tingkat kesalahan estimasi kemampuan peserta Tes sekecil mungkin.

2.4 Rencana Pengembangan CAT

Alur yang akan diterapkan pada sistem ini memakai aturan yang sama dengan konsep dari CAT. Sistem dibuat untuk melakukan penilaian secara digital sehingga kebutuhan sistem disesuaikan dengan penggunanya. Hal ini diharapkan akan memperluas pengetahuan tentang suatu sistem penilaian menggunakan komputer khususnya dengan menggunakan *Computer Adaptive Test*. Hal-hal yang diharapkan bisa dibuat dalam sistem ini adalah:

- a. Konsep CAT dalam penilaian hasil belajar, dengan memperhatikan konsep
 CAT di terapkan dalam salah satu bahasa pemprograman sehingga dapat
 diterapkan dalam rujukan sistem penilaian.
- b. Sistem penilaian hasil belajar siswa, sistem mudah dipahami oleh user. Dan mudah untuk dipahami sebagai suatu alat bantu dalam penilaian.

Diharapkan dari rancangan yang sudah direncanakan akan dapat digunakan sebagai inovasi suatu sistem penilaian dalam perkembangan teknologi.

2.5 HYPERTEXT PREPROCESSOR (PHP)

Menurut Didik Dwi Presetyo (2004: 76), PHP (*Hypertext Preprocessor*) merupakan bahasa *scripting server-side*, dimana pemrosesan datanya dilakukan pada sisi server. Sederhananya, serverlah yang akan menerjemahkan *script* program, baru kemudian hasilnya akan dikirim kepada *client* yang melakukan permintaan.

PHP (Hypertext Prepocessor) adalah bahasa skrip yang dapat ditanamkan atau disisipkan ke dalam HTML (Hyper Text Markup Language). PHP banyak dipakai untuk memrogram situs web dinamis. Seluruh aplikasi berbasis web dapat dibuat dengan PHP. Namun kekuatan yang paling utama PHP adalah pada konektivitasnya dengan sistem database di dalam web. Kelebihan-kelebihan dari PHP diantaranya adalah:

a. PHP mudah dibuat dan dijalankan, maksudnya PHP dapat berjalan dalam Web Server dan dalam Sistem Operasi yang berbeda pula.

- b. PHP adalah *software open-source* yang gratis dan bebas didistribusikan kembali di bawah lisensi GPL (GNU *Public License*). User dapat men-download kode-kode PHP tanpa harus mengeluarkan uang atau khawatir dituntut oleh pihak pencipta PHP.
- c. PHP bisa dioperasikan pada *platform* Linux ataupun Windows.
- d. PHP sangat efisien, karena PHP hanya memerlukan *resource system* yang sangat sedikit dibanding dengan bahasa pemograman lain.
- e. Ada banyak Web Server yang mendukung PHP, seperti Apache, PWS, IIS, dan lain-lain.
- f. PHP juga didukung oleh banyak *database*, seperti MySQL, PostgreSQL, Interbase, SQL, dan lain-lain.
- g. Bahasa pemograman PHP sintaknya sederhana, singkat dan mudah untuk dipahami.
- h. HTML *embedded*, artinya PHP adalah bahasa yang dapat ditulis dengan menempelkan pada sintak sintak HTML.

2.6 MySQL

MySQL merupakan salah satu jenis database server. MySQL termasuk database server yang banyak digunakan dan sangat terkenal. Hal ini disebabkan bersifat *free* (tidak perlu membayar untuk menggunakannya). Selain itu MySQL menggunakan SQL sebagai Bahasa dasar untuk mengakses database yang terbagi

UNIVERSITAS NEGERI SEMARANG

menjadi tiga bagian, yaitu DDL, DML dan DCL. Menurut Samsul Hadi (2013: 51) dijelaskan yaitu:

A. DDL

DDL (Data Definition Language) merupakan kelompok perintah yang berfungsi untuk mendefinisikan atribut-atribut database, tabel, atribut (kolom), batasan-batasan terhadap suatu atribut, serta hubungan antar tabel. Perintah yang termasuk dalam kelompok DDL adalah:

1. CREATE untuk membuat tabel

Contoh: CREATE TABLE peserta (nomor_peserta INTERGER, nama_peserta VARCHAR(45));

2. ALTER untuk mengubah struktur tabel

Contoh: ALTER TABLE peserta ADD password VARCHAR(30);

3. DROP untuk menghapus tabel.

Contoh: DROP TABLE peserta;

B. DML

DML (*Data Manipulation Language*) adalah perintah yang berfungsi untuk memanipulasi data dalam database. Misalnya untuk pengambilan, penyisipan, pengubahan dan penghapusan data. Perintah yang masuk kategori DML adalah:

1. SELECT untuk memilih data, contoh: SELECY *FROM peserta.

- 2. INSERT untuk menambah data. Contoh : INSERT INTO peserta VALUES ('1', 'khoir', 'rahasia');
- 3. DELETE untuk menghapus data. Contoh: DELETE FROM siswa WHERE nama='hakim';
- 4. UPDATE untuk mengubah data. Contoh: UPDATE peserta SET nama:'dina' where nomor peserta='1';

C. DCL

DCL (*Data Control Language*) Berisi perintah untuk mengendalikan pengaksesan data. Pengendalian dilakukan berdasarkan perpengguna, per-tabel, per-kolom maupun per-operasi yang boleh dilakukan. Perintah-perintah yang termasuk dalam DCL adalah sebagai berikut:

- 1. GRANT untuk memberikan kendali pengaksesan data. Contoh :GRANT SELECT peserta TO Khoir;
- 2. REVOKE untuk mencabut kemampuan pengaksesan data. Contoh REVOKE SELECT ON peserta FROM khoir;
- 3. LOCK TABLE untuk mengunci tabel. Contoh: LOCK TABLE peserta;

2.7 Penelitian Yang Relevan

Penelitian tentang implementasi Computer Adaptive Test sebagai sistem penilaian hasil belajar siswa ini mempuanyai acuan ataupun referensi dari penelitian

yang telah dilakukan oleh para penelitian sebelumnya, judul penelitian tersebut adalah:

- 1. Penelitian yang dilakukan oleh Lalang Erawan dan Stefanus Santoso. Jurnal berjudul: "Computer Adaptive Test dengan Pendekatan Item Response Theory (IRT) Satu Parameter". Hasil penelitian: sistem dikembangkan berbasis web dan dengan pendekatan model IRT satu parameter yaitu tingkat kesukaran soal. Hasil penelitian menunjukan bahwa responden memberikan penilaian sangat baik. Sistem evaluasi yang dikembangkan lebih akurat dalam mengukur kemampuan pengguna dan dapat mengakomodasi keberagaman kemampuan yang dimiliki pengguna.
- 2. Penelitian yang dilakukan oleh Rukli dan S. Hartati. Jurnal berjudul:
 "Penerapan Sistim Pendukung Keputusan dalam Sistem Pengujian Computer
 Adaptive Testing". Hasil penelitian: sistem dikembangkan berbasis web dan
 dijadikan sebagai sistem pendukung keputusan berdasarkan model logistic
 satu parameter dengan memperhatikan parameter tingkat kesulitan soal yang
 digunakan untuk proses seleksi ujian masuk. Soal yang direspon oleh
 responden hanya berjumlah 10 soal untuk setiap kali tes sampai ditemukan
 tingkat kemampuannya.
- 3. Penelitian yang dilakukan oleh Agus Santoso. Jurnal berjudul: "Pengembangan *Computerized Adaptive Testing* untuk Mengukur Hasil Belajar Mahasiswa Universitas Terbuka". Hasil penelitian: bank soal data

empiris akurat untuk mengukur peserta tes dengan kemampuan sedang sampai tinggi (-1,2 sampe +3,0) namun kurang akurat dalam mengukur kemampuan rendah (-3 sampai -1,5) banyaknya butir soal yang diperlukan pada desain CAT murni berkisar antara 8 sampai 15 butir dalam sekali tes.

2.8 Kerangka Berpikir

Berbagai usaha telah dilakukan para ahli pengukuran pendidikan untuk mengatasi masalah dalam penilaian hasil belajar. Prosedur penyusunan soal kriteria soal yang baik yang baik telah dikembangkan. Kemajuan teknologi komputer dan teori pengujian telah mempercepat perubahan format ujian dari konvensional ke pengujian berbasis komputer. Dengan memanfaatkan komputer dalam penilaian hasil belajar, maka kerjasama dan kecurangan yang dilakukan peserta tes dapat dikurangi; selain itu pemberian skor atau nilai kepada peserta tes akan objektif dan konsisten.

Salah satu ujian berbasis komputer adalah computer adaptive test, suatu konsep tes berbasis komputer dimana butir soal (tes) yang diberikan disesuaikan dengan tingkat kemampuan setiap peserta tes. Keuntungan menggunakan computer adaptive test, waktu yang diperlukan untuk ujian lebih singkat, peserta didik dalam mengerjakan soal bersifat individual. Hal ini berarti bahwa butir yang disajikan untuk tiap peserta didik berbeda, sehingga mengurangi peluang bekerjasama.

Berdasarkan kerangka berfikir diatas dapat diambil asumsi bahwa *computer* adaptive test dapat digunakan sebagai sistem penilaian hasil belajar.

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan penelitian yang telah dilakukan, didapatkan kesimpulan penelitian sebagai berikut:

- 1. Pengembangan *computer adaptive test(CAT)* sebagai sistem penilaian hasil belajar ini dikembangkan berbasis website dengan menggunakan bahasa pemprograman PHP dan *database* MySQL. Penelitian rancang bangun dengan menggunakan metode pengembangan *Waterfall* dan pengujian black-box dan white-box serta uji tanggapan menggunakan kuisioner.
- 2. Pada uji coba respon pengguna sistem penilaian hasil belajar menggunakan *Computer Adaptive Test* (CAT) berada pada kategori baik dengan nilai rata-rata 81,25%.

5.2 Saran

Pengembangan aplikasi yang dilakukan masih memiliki banyak kekurangan, untuk itu saran masih dibutuhkan dalam penelitian ini. Berikut merupakan saran untuk pengembangan penelitian:

UNIVERSITAS NEGERI SEMARANG

- Sistem penilaian hasil belajar menggunakan computer adaptive test sebagai alat bantu penilaian masih sebatas dalam 1 mata pelajaran, sehingga bias dikembangkan dalam database yang lebih luas lagi sehingga dapat digunakan dalam skala luas.
- 2. Sistem penilaian hasil belajar menggunakan *computer adaptive test* ini dapat dikembangkan dengan menggunakan parameter yang lebih banyak sehingga hasil yang didapatkan menjadi lebih akurat.
- 3. Sistem penilaian hasil belajar menggunakan *computer adaptive test* dalam penelitian ini dapat dikembangkan ke dalam Bahasa pemrograman lain sehingga dapat digunakan tidak hanya sebatas desktop saja.



DAFTAR PUSTAKA

- Ali, Muh. 1982. Penelitian Kependidikan Prosedur dan Strategi. Bandung: Angkasa
- Arikunto, Suharsimi. 2006. *Prosedur Penelitian Suatu Pendekatan Praktik*. Jakarta: Rineka Cipta
- Arikunto, Suharsimi. 2007. Dasar-Dasar Evaluasi Pendidikan. Jakarta: Bumi Aksara
- Sugiono. 2014. Metode Penelitian Pendidikan (Pendekatan Kuantitatif, Kualitatif, dan R&D). Bandung: Alfabeta.
- Arifin, Zaenal. 2014. Evaluasi Pembelajaran. Bandung: Remaja Rosdakarya.
- Hadi, Samsul. 2013. Pengembangan Computerized Adaptive Test Berbasis Web.

 Yogyakarta: Aswaja Presindo.
- Raharjo, Budi. 2011. Belajar Pemrograman Web. Bandung: Modula.
- Daryanto. 2010. *Media Pembelajaran*. Yogyakarta: Gava Media.
- Fatkhudin, A dkk. 2014. *Item Response Theory Model Empat Parameter Logistik*Pada Computerized Adaptive Test, Jurnal Sistem Informasi Bisnis 02(2014)
- Hambleton, R.K., 1993. *Principles and Selected Applications of Item Response Theory*. In Linn, Robert L. (Editor). Educational Measurement. Third Edition. Phoenix: American Council on Education, Series on Higher Education Oryx Press.
- Rukli dan S Hartati. 2011. Penerapan Sistem Pendukung Keputusan dalam Sistem Pengujian Computerized Adaptive Testing. IJCCS, Volume 5 No. 3, November 2011.

- Suwarto. 2009. Computerized Adaptive Testing (CAT). Widyatama 141-144.
- Erawan, L dan Stefanus S. 2009. *Computer-Adaptive Test dengan Pendekatan Item**Response Theory Satu Parameter. Jurnal Teknologi Informasi, Volume 5

 Nomor 2, Oktober 2009, ISSN 1414-9999.
- Kamaruddin dan Haryanto. 2014. Pengembangan Sistem Penilaian Hasil Belajar Mata Pelajaran Menganalisis Rangkaian Listrik Berbasis Computerized Adaptive Testing. Jurnal Pendidikan Vokasi, Vol 4, Nomor 1, Februari 2014.
- Santoso, Agus. 2010. Pengembangan Computerized Adaptive Testing Untuk

 Mengukur Hasil Belajar Mahasiswa Universitas Terbuka. Jurnal Penelitian

 dan Evaluasi Pendidikan Tahun 14, Nomor 1, 2010.
- Pressman, RS. 2002. Rek<mark>ayasa Perangkat Lunak Pe</mark>ndekatan Praktisi (**Buku Satu**).

 Yogyakarta: Andi.
- Kusaeri dan Suprananto. 2012. *Pengukuran dan Penilaian Pendidikan*. Yogyakarta: Graha Ilmu.

UNIVERSITAS NEGERI SEMARANG