PERBANDINGAN PENERAPAN ALGORITMA TEORI RESPON BUTIR DAN FUZZY TSUKAMOTO PADA *COMPUTERIZED ADAPTIVE TEST*

Wahyuni¹, Muhammad Fahmi²

^{1,2}STMIK Widya Cipta Dharma, Samarinda

Email : wahyuni@wicida.ac.id⁽¹⁾, mfahmi@wicida.ac.id⁽²⁾

Abstrak - Pada dasarnya sistem CAT (*Computerized Adaptive Test*) hampir sama dengan sistem CBT (*Computerized Based Testing*) dimana test dilakukan melalui komputer, namun sistem CAT dapat merandom dan memberikan soal sesuai dengan kemampuan pesertanya. Jika peserta tes tersebut memiliki kemampuan rendah, maka sistem akan memberikan soal yang cenderung mudah. Sebaliknya, jika peserta tes memiliki kemampuan tinggi, maka sistem akan memberikan soal yang cenderung sulit. Penelitian ini bertujuan untuk mengetahui metode manakah yang paling tepat untuk diterapkan pada sistem *Computerized Adaptive Test* dalam pemberian butir soal kepada peserta tes sehingga menghasilkan hasil yang maksimal. Pemilihan butir soal dimulai dari butir soal kedua sampai soal yang terakhir. Metode Teori Respon Butir atau *Item Response Theory* (IRT) pada penelitian ini menggunakan tiga parameter logistik (3PL). Metode analisis yang dilakukan berupa pengujian kepada responden secara manual, menggunakan sistem CAT dengan metode IRT dan menggunakan sistem CAT dengan metode Fuzzy Tsukamoto. Setelah dilakukan pengujian, maka dibandingkan hasil yang telah didapat dari masing-masing pengujian tersebut untuk mengetahui metode manakah yang paling baik untuk diterapkan pada sistem CAT ini. Hasil yang di dapat adalah metode IRT dan Fuzzy sama-sama dapat diterapkan dalam sistem CAT, namun metode IRT lah yang lebih baik untuk diterapkan dibandingkan dengan metode Fuzzy Tsukamoto.

Kata Kunci – *Computerized Adaptive Test, Teori Respon Butir, Fuzzy Tsukamoto*

1. PENDAHULUAN

Computerized Adaptive Test (CAT) adalah sistem pengujian berbantuan komputer yang lebih maju dibanding *Computerized Base Testing*[1]. Tidak seperti CBT, komputer hanya dimanfaatkan untuk memindahkan tes dari form kertas ke monitor komputer, pada CAT komputer diatur dan dimanfaatkan sedemikian rupa hingga dapat memilih dan menampilkan butir-butir soal yang sesuai dengan kemampuan individu peserta tes. Karena tes disesuaikan dengan kemampuan peserta tes, maka tak ada pertanyaan yang lebih sukar ataupun lebih mudah. Peserta yang berkemampuan tinggi hanya akan diberikan butir soal dengan tingkat kesukaran tinggi, begitupun sebaliknya, bagi peserta yang berkemampuan rendah hanya akan diberi butir-butir soal dengan tingkat kesukaran rendah dan tidak perlu diberikan butir soal yang sukar. Karena CAT hanya menyajikan soal yang tingkat kesukarannya sesuai dengan kemampuan peserta tes, maka jumlah butir soal yang disajikan oleh CAT bisa lebih sedikit dibandingkan dengan soal yang disajikan dalam bentuk *Paper Pencil Test* (PPT) dengan ketelitian pengukuran yang sama bahkan lebih baik. Dengan

demikian, penggunaan CAT dapat mengurangi jumlah waktu dan biaya untuk pengadministrasian tes dan penyusunan butir-butir soal dalam bank soal.

Teori respon butir atau *Item Response Theory* (IRT) menggambarkan peluang menjawab butir soal secara benar berdasarkan tingkat kemampuan peserta tes dan butir soal yang diberikan[1]. Umumnya ada tiga model logistik dalam teori respons butir, yaitu model satu parameter logistik (1PL), model dua parameter logistik (2PL) dan model tiga parameter logistik (3PL). Perbedaan dari ketiga model ini terletak pada banyaknya parameter yang digunakan dalam menggambarkan karakteristik butir soal dalam model yang digunakan. Parameter-parameter yang digunakan tersebut adalah indeks daya beda butir (disimbolkan dengan a), indeks kesukaran (disimbolkan dengan b) dan indeks tebakan semu (disimbolkan dengan c). Sedangkan kemampuan peserta tes dilambangkan dengan θ .

Sebuah penelitian telah mencoba untuk menambahkan parameter baru untuk metode IRT sehingga menjadi model empat parameter logistik (4PL) yaitu *carelessness*, lalu membandingkan IRT 3PL dengan IRT model 4P untuk penilaian

menggunakan *computerized adaptive test*[2]. Hasil dari penelitian tersebut adalah algoritma model 4PL dapat mengukur kemampuan peserta tes lebih singkat dan cepat, dan peneliti menegaskan peluang peserta tes menjawab benar cenderung naik. Namun pada penelitian ini, tidak dijelaskan apakah hasil pengukuran atau estimasi kemampuan peserta menjadi lebih baik atau tidak. Penelitian selanjutnya adalah mengembangkan *computerized adaptive test* menggunakan algoritma fuzzy dengan metode Tsukamoto berbasis web[3]. Pada penelitian ini peneliti mencoba menggantikan algoritma teori respon butir dengan algoritma fuzzy menggunakan metode Tsukamoto dalam pengambilan keputusan untuk pemberian butir soal. Hasil dari penelitian ini mengatakan bahwa algoritma fuzzy dengan metode Tsukamoto dapat diterapkan dalam pengambilan keputusan untuk memilih soal pada sistem CAT, namun tidak dijelaskan apakah hasil tersebut dapat berjalan lebih baik dari pada penerapan algoritma teori respon butir yang umum digunakan atau tidak. Selain itu sebelumnya telah dikembangkan sistem *computerized adaptive test* menggunakan teori respon butir namun hanya menggunakan model 1PL[4]. Dimana tentunya akan lebih baik menggunakan model 3PL sehingga didapatkan hasil yang lebih akurat.

Pada penelitian ini akan diterapkan sistem *computerized adaptive test* pada tes online yang akan menggunakan algoritma teori respon butir model 3PL dalam perhitungan untuk mengestimasi kemampuan peserta tes. Berdasarkan uraian di atas, tujuan dari penelitian ini adalah mencari algoritma manakah yang paling tepat untuk diterapkan pada program CAT dalam melakukan pemilihan butir soal kedua dan soal selanjutnya sampai tes selesai.

Melalui sistem yang dibuat ini diharapkan dapat lebih membantu para penguji untuk mengukur kemampuan masing-masing individu peserta tes dengan tidak hanya melihat dari nilai yang dicapai seperti yang dilakukan pada PPT, namun juga dapat menganalisis titik kelemahan peserta tes melalui butir-butir soal yang diberikan oleh sistem. Selain itu sistem ini juga diharapkan dapat meminimalisir waktu pelaksanaan tes, dan meminimalisir waktu dan biaya pengadministrasian dalam mempersiapkan ujian atau tes tersebut.

2. METODE

2.1 *Computerized Adaptive Test*

Computerized Adaptive Test (CAT) merupakan generasi kedua dari penggunaan komputer untuk pengujian [1]. *Adaptive* berarti bahwa butir soal yang diberikan disesuaikan dengan tingkat kemampuan setiap peserta tes atau *tailored testing* [1]. Menurut *National Council on Measurement in Education*, *adaptive test* merupakan tes berbasis komputer dimana butir soal yang akan di berikan kepada peserta tes bergantung benarnya respon peserta tes pada soal yang paling akhir atau baru saja dikerjakan [5]. CAT memberikan soal kepada peserta tes dengan tujuan memaksimalkan tingkat keakuratan tes bergantung pada respon terhadap soal yang diberikan sebelumnya. Pada CAT komputer tidak hanya sekedar memindahkan butir soal ke dalam komputer, tetapi komputer diatur untuk menyeleksi dan menyajikan butir soal menurut perkiraan tingkat kemampuan peserta tes. Hal ini mengakibatkan individu peserta tes yang memiliki tingkat kemampuan tinggi akan mendapatkan butir soal yang lebih sulit dibandingkan dengan individu yang memiliki tingkat kemampuan rendah. Sebaliknya individu peserta tes yang memiliki tingkat kemampuan rendah akan mendapatkan butir soal yang lebih mudah dibandingkan dengan peserta tes yang memiliki tingkat kemampuan tinggi. Pertama peserta akan diberi soal dengan tingkat kesulitan rata-rata atau sedang (rata-rata atau sedang = 0). Jika peserta menjawab soal dengan benar maka sistem akan mencari ke dalam database dan mengambil butir soal yang tingkat kesulitannya sedikit di atas atau sedikit lebih sulit dari tingkat kesulitan sebelumnya. Sebaliknya, jika peserta menjawab soal dengan jawaban yang salah, maka sistem akan mencari ke dalam database dan mengambil soal dengan tingkat kesulitan sedikit di bawah atau sedikit lebih mudah dari soal sebelumnya. Dengan demikian CAT lebih efisien karena dapat mengestimasi kemampuan peserta tes dengan jumlah butir soal yang lebih sedikit dibandingkan PPT atau CBT tanpa mengurangi ketepatan pengukuran.

Prinsip *Computerized Adaptive Test (CAT)* adalah sebagai berikut [6]: (1) Membangun bank soal, bank soal umumnya menggunakan teori respon

butir. (2) Prosedur pemilihan item awal (starting rule), biasanya CAT dapat dimulai dengan memilih butir soal dengan tingkat kesukaran sedang. (3) Prosedur pemilihan item selama tes berjalan. (4) Prosedur untuk mengakhiri tes (stopping rule), biasanya tes CAT berhenti apabila item pada bank soal telah habis dan tingkat ketelitian estimasi kemampuan peserta telah tercapai, maksudnya jika soal yang disajikan belum habis, namun hasil estimasi kemampuan peserta tes telah konsisten 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 dicapainya. (5) Mengestimasi kemampuan siswa.

Berdasarkan cara kerja CAT di atas, maka dapat digambarkan cara kerja dari CAT seperti pada Gambar 1 di bawah ini:

Gambar 1. Cara Kerja CAT

2.2 Algoritma CAT dengan Teori Respon Butir

Prinsip kerja dari algoritma ini adalah sebagai berikut[1]. Mula-mula peserta tes diberi soal dengan tingkat kesulitan awal yang sedang ($b_{awal}=0$) karena dianggap mempunyai tingkat kemampuan awalnya (θ_{awal}) = sedang = 0. Peserta tes diberi kesempatan untuk menjawab soal dengan alokasi waktu tertentu. Jika soal dengan tingkat kesulitan sedang tersebut dijawab dengan benar, peserta diberi soal baru yang lebih sulit, jika dijawab salah maka peserta diberi soal yang lebih mudah. Kemudian kemampuan (θ) setelah menjawab soal baru $P(\theta)$, Q

(θ), $I_i(\theta)$, $SE(\theta)$ dan harga mutlak selisih kesalahan baku antar penyajian soal dihitung. Rumus yang digunakan untuk menghitung kemampuan (θ), probabilitas menjawab benar berdasarkan kemampuan ($P(\theta)$), probabilitas menjawab salah ($Q(\theta)$), fungsi informasi butir ($I_i(\theta)$), dan kesalahan baku ($SE(\theta)$) adalah sebagai berikut:

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

$$P_i(\theta) = c_i + \frac{(1-c_i)e^{Da_i(\theta-b_i)}}{1 + e^{Da_i(\theta-b_i)}} \quad (2)$$

$$Q_i(\theta) = 1 - P_i(\theta) \quad (3)$$

$$I_i(\theta) = \frac{[P_i(\theta)]^2}{P_i(\theta)Q_i(\theta)} \quad (4)$$

$$SE(\theta) = \frac{1}{\sum_{i=1}^N I_i(\theta)} \quad (5)$$

2.3 Algoritma CAT dengan Fuzzy Tsukamoto

Algoritma ini digunakan untuk inferensi pengambilan keputusan butir soal [7]. Pada metode Tsukamoto setiap konsekuensi pada aturan yang berbentuk *IF-Then* direpresentasikan dengan suatu himpunan fuzzy dengan fungsi keanggotaan yang monoton. Sebagai hasilnya, output hasil inferensi dari tiap-tiap aturan diberikan secara tegas (crisp) berdasarkan α -predikat (fire strength). Hasil akhirnya diperoleh dengan menggunakan rata-rata terbobot. Pembentukan himpunan fuzzy (fuzzyfikasi) adalah sebagai berikut: (1)Tingkat kesulitan soal (b) dikatakan baik jika nilai berkisar antara -2 sampai +2. (2) Daya beda (a), indeks daya pembeda butir tidak bersifat negatif sehingga range daya beda adalah 0 - 2. (3) Tebakan semu (c) atau kebetulan menjawab benar memiliki range 0 - 0.25. (4)Tingkat kemampuan (θ) yang digunakan adalah -1 sampai +1. Dalam sistem ini semua himpunan fuzzy nya menggunakan kurva linier yaitu kurva naik dan turun. Rumus untuk mencari nilai keanggotaan adalah:

$$\mu_{turun} = \frac{(\text{NilaiMaksHimpunan} - \text{Nilaidiketahui})}{\text{Range}} \quad (6)$$

$$\mu_{naik} = \frac{(\text{Nilaidiketahui} - \text{NilaiMinHimpunan})}{\text{Range}} \quad (7)$$

$$\mu = \text{nilai keanggotaan fuzzy.} \quad (8)$$

Selanjutnya melakukan inferensi dari beberapa rule seperti tabel 1 di bawah ini.

Tabel 1. Aturan algoritma fuzzy

Rule	Kesulitan soal (b)	Daya Beda (a)	Tebakan Semu (c)	Tingkat Kemampuan (θ)
R1	Naik	Naik	Naik	Naik
R2	Naik	Naik	Turun	Naik
R3	Naik	Turun	Naik	Naik
R4	Naik	Turun	Turun	Naik
R5	Turun	Naik	Naik	Turun
R6	Turun	Naik	Turun	Turun
R7	Turun	Turun	Naik	Turun
R8	Turun	Turun	Turun	Turun

Sedangkan untuk mencari nilai α predikat dapat digunakan rumus sebagai berikut: α predikat = minimum(μ_1, μ_2). Untuk mencari nilai z (daerah kodomain) adalah sama seperti rumus untuk mencari nilai keanggotaan sesuai dengan kondisi rule nya.

Selanjutnya dilakukan defuzzifikasi yaitu dengan menghitung rerata bilangan kodomain yang dihasilkan dengan menggunakan rumus Tsukamoto sebagai berikut:

$$z = (a_1 * z_1 + a_2 * z_2 + a_3 * z_3 + a_4 * z_4) / (a_1 + a_2 + a_3 + a_4) \quad (9)$$

Untuk daya beda, tingkat kesulitan dan tebakan semu yang dijadikan perhitungan adalah nilai daya beda, tingkat kesulitan dan tebakan semu soal yang dijawab benar terakhir kali. Hasil dari perhitungan ini adalah nilai tingkat kemampuan siswa yang kemudian menjadi patokan untuk mencari butir soal selanjutnya.

2.4 Alur Penelitian

Adapun alur penelitian pada penelitian ini adalah sebagai berikut:

1. Studi literature, pada penelitian ini tahap studi literature berupa mencari dan mempelajari literature pustaka dari sejumlah jurnal dan penelitian serta buku-buku yang berkaitan dengan Teori Respon Butir, *Fuzzy Logic* dan *Computerized Adaptive Test*.
2. Observasi, selanjutnya peneliti mengobservasi soal-soal yang layak untuk dijadikan item butir pada sistem CAT.
3. Perancangan, membuat diagram arsitektur sistem dan alur program berbentuk flowchart. Setelah dibuat alur program maka masuk pada tahap perancangan database. Untuk arsitektur sistem dibagi menjadi dua bagian yaitu bagian admin dan bagian peserta tes. Halaman admin berisi halaman untuk menginputkan butir tes ke dalam bank soal pada sistem CAT, dan pengelolaan data yang lain

seperti data peserta tes dan hasil dari peserta tes. Sedangkan di halaman peserta terdapat halaman tes untuk peserta. Sistem ini akan menggunakan tiga buah variable dan menerapkan model teori respon butir tiga parameter. *Stopping rule* yang digunakan pada sistem ini adalah jika butir soal habis dan jika tingkat ketelitian estimasi kemampuan telah tercapai, sedangkan waktu pengerjaan soal tidak dijadikan *stopping rule*.

4. Implementasi, pada tahap ini sistem yang akan dibangun adalah sistem *Computerized Adaptive Test* berbasis web. Sedangkan untuk basis datanya menggunakan MySQL. Pada penelitian ini akan dibuat dua buah modul dimana modul yang pertama adalah sistem CAT menggunakan algoritma teori respon butir dalam pencarian butir selanjutnya, dan modul kedua menggunakan algoritma fuzzy Tsukamoto. Sistem ini menggunakan satu buah database yang sama.

5. Pengujian, pada tahap ini kedua sistem akan diujikan kepada sepuluh orang responden sebagai peserta tes. Selain melakukan pengujian terhadap sistem CAT, peserta tes juga akan mengerjakan soal secara manual. Setelah melakukan pengujian maka dilakukan proses membandingkan hasil dari kedua sistem CAT tersebut terhadap hasil peserta tes secara manual. Dari hasil yang di dapatkan pada proses perbandingan, maka dapat disimpulkan algoritma mana yang hasilnya paling mendekati dengan kemampuan peserta sebenarnya. Selain itu dapat pula disimpulkan kelebihan dan kelemahan dari masing-masing algoritma yang digunakan.

3. HASIL DAN PEMBAHASAN

3.1 Hasil Observasi

Peneliti telah melakukan observasi terhadap soal-soal TOEFL yang layak untuk dijadikan item butir pada sistem CAT. Dari hasil observasi tersebut, peneliti mendapatkan 114 butir soal dari 160 butir soal yang dapat diberikan kepada peserta tes. Tingkat kesulitan soal (b) dikatakan baik jika nilai berkisar antara -2 sampai +2. Daya beda (a), indeks daya pembeda butir tidak bersifat negatif sehingga range daya beda adalah 0 - 1. Tebakan semu (c) atau kebetulan menjawab benar memiliki range 0 - 0.25 jika pilihan soal pilihan ganda terdapat empat buah pilihan. Tingkat kemampuan (θ) yang digunakan adalah -1 sampai +1. Adapun rentang

parameter yang didapat dari 114 butir soal setelah dilakukan observasi dapat dilihat pada tabel 2.

Tabel 2. Rentang Soal Yang Didapat

Parameter	Max	Min
a	1	0.18
b	1.98	-1.80
c	0.25	0.03

3.2 Hasil Perancangan

Sistem ini memiliki dua buah level user yaitu admin dan peserta tes. Arsitektur sistem untuk kedua level user tersebut dapat dilihat pada gambar 2 di bawah ini.

Gambar 2. Arsitektur Sistem

3.2 Implementasi

Sistem CAT ini dibuat berbasis website menggunakan bahasa pemrograman PHP, HTML dan javascript. Adapun tampilan dari sistem ini adalah sebagai berikut.

Gambar 3. Halaman Awal Sistem.

Untuk melihat detail hasil tes dari peserta, maka kita bisa melihatnya melalui halaman admin. Pada gambar 4, dapat kita lihat bahwa sistem akan

memberikan soal pertama yang tingkat kesukarannya (b) sedang yaitu 0. Setelah menjawab respon dari peserta tes, maka sistem akan menghitung θ setelah menjawab soal baru, $P(\theta)$, $Q(\theta)$, $I_i(\theta)$, $SE(\theta)$ dan harga mutlak selisih kesalahan baku antar penyajian soal. Jika respon ppeserta tes untuk soal tersebut benar, maka diberikan nilai 1. Jika salah maka diberikan nilai 0. Jika nilai respon peserta = 1, maka θ setelah menjawab soal baru akan dihitung sesuai dengan persamaan 1, tetapi jika respon jawaban peserta bernilai 0, maka θ setelah menjawab soal baru = θ sebelum menjawab soal baru. Pada gambar 4, dapat dilihat bahwa peserta tes mendapatkan soal sebanyak 19 buah. Banyaknya soal yang diberikan oleh sistem bergantung pada respon dari peserta itu sendiri. Sistem akan berhenti memberikan soal atau tercapainya *stopping rule* adalah jika selisih kesalahan baku sudah mencapai 0.01 atau kurang dari 0.01 atau jika soal yang diberikan sudah habis. Tampilan sistem setelah memilih metode adalah sama, walaupun peserta menggunakan metode IRT maupun metode fuzzy. Hanya butir soal yang diberikan kepada peserta saja yang berbeda yang mana perbedaan pemberian butir soal tersebut mempengaruhi nilai akhir yang di dapat oleh peserta.

CAT														
Detail Hasil Tes														
No.	i	b	a	c	Skor	θ Awal	θ Sd Jwb	$P(\theta)$	$Q(\theta)$	IIF	$SE(\theta)$	Selisih SE		
1	27	0	0.91	0	1	0	0	0.5	0.5	0.25	2	2		
2	68	0.1	0.78	0.1	0	0	0	0.48	0.52	0.25	1.414	0.586		
3	62	0	0.89	0.06	1	0	0.054	0.549	0.451	0.248	1.156	0.258		
4	82	0.1	1	0.06	1	0.054	0.16	0.554	0.446	0.247	1.003	0.153		
5	76	0.2	1	0.09	0	0.16	0.16	0.47	0.53	0.249	0.897	0.106		
6	5	0.1	0.82	0.08	0	0.16	0.16	0.441	0.559	0.247	0.819	0.078		
7	35	0	0.45	0.18	0	0.16	0.16	0.385	0.615	0.237	0.761	0.058		
8	77	-0.1	0.78	0.05	1	0.16	-0.06	0.538	0.462	0.249	0.711	0.05		
9	48	0	0.73	0.19	0	-0.06	-0.06	0.42	0.58	0.244	0.671	0.04		
10	18	-0.1	0.82	0	0	-0.06	-0.06	0.486	0.514	0.25	0.636	0.035		
11	7	-0.2	0.55	0.1	1	-0.06	-0.149	0.561	0.439	0.246	0.607	0.029		
12	1	-0.1	0.55	0.07	1	-0.149	-0.062	0.543	0.457	0.248	0.581	0.026		
13	60	0.1	0.78	0.1	1	-0.062	0.172	0.571	0.429	0.245	0.558	0.023		
14	83	0.2	1	0.08	1	0.172	0.277	0.57	0.43	0.245	0.538	0.02		
15	8	0.3	0.36	0.15	0	0.277	0.277	0.428	0.572	0.245	0.52	0.018		
16	23	0.2	0.27	0	0	0.277	0.277	0.491	0.509	0.25	0.503	0.017		
17	30	0.1	0.73	0	0	0.277	0.277	0.445	0.555	0.247	0.488	0.015		
18	55	-0.3	0.45	0.11	0	0.277	0.277	0.348	0.652	0.227	0.475	0.013		
19	45	-0.41	0.73	0.12	0	0.277	0.277	0.263	0.737	0.194	0.465	0.01		
θ Akhir					0.277									
Skor					54.62									

Gambar 4. Halaman Detail Hasil Tes

3.3 Pengujian Secara Manual

Melihat hasil pengujian yang dilakukan secara manual, maka dapat digambarkan melalui grafik di bawah ini.

Gambar 4. Grafik Persentase Kemampuan Responden Secara Manual

Berdasarkan gambar 4 di atas dapat dilihat bahwa responden yang berkemampuan tinggi sebanyak 23%, berkemampuan rendah 10% dan berkemampuan sedang sebanyak 67%. Artinya rata-rata responden berkemampuan sedang.

3.4 Pengujian Metode Teori Respon Butir (IRT)

Tingkat kesesuaian antara pengujian secara manual dengan yang menggunakan metode IRT dapat dilihat pada gambar 5 di bawah ini.

Gambar 6. Grafik Kesesuaian Pengujian Manual dengan Pengujian IRT

Berdasarkan pada gambar 6 di atas dapat dilihat bahwa 70% dari hasil pengujian manual dan metode IRT saling berkesesuaian atau bisa dikatakan bahwa hasil yang didapatkan berada pada range yang sama. Sedangkan 30% dari data tidak berkesesuaian dimana terdapat perbedaan antara hasil yang didapatkan secara manual dengan hasil yang menggunakan metode IRT.

3.5 Pengujian metode Fuzzy Tsukamoto

Tingkat kesesuaian antara pengujian secara manual dengan yang menggunakan metode fuzzy dapat dilihat pada gambar 7 di bawah ini.

Gambar 8. Grafik Kesesuaian Pengujian Manual dengan Pengujian Fuzzy

Berdasarkan pada gambar 8 di atas dapat dilihat bahwa 80% dari hasil pengujian manual dan metode fuzzy saling berkesesuaian atau bisa dikatakan bahwa hasil yang didapatkan berada pada range yang sama. Sedangkan 20% dari data tidak berkesesuaian dimana terdapat perbedaan antara hasil yang didapatkan secara manual dengan hasil yang menggunakan metode fuzzy.

3.6 Perbandingan Pengujian Metode IRT dan Metode Fuzzy Tsukamoto

Berikut adalah grafik kenaikan dan penurunan θ akhir yang dihasilkan oleh kedua metode tersebut dengan mengambil 10 data responden.

Gambar 9. Grafik Kenaikan dan Penurunan θ Akhir Metode IRT

Gambar 10. Grafik Kenaikan dan Penurunan θ Akhir Metode Fuzzy Tsukamoto

Berdasarkan pada gambar 9 dan gambar 10 di atas dapat dilihat bahwa nilai θ akhir yang dihasilkan dari metode IRT memiliki range yang lebih luas dan bervariasi yaitu mulai dari -0.06 sampai 0.82 dan tujuh orang mendapatkan nilai dalam range sedang, tiga orang mendapatkan nilai dalam range tinggi, sedangkan θ akhir yang dihasilkan melalui metode fuzzy memiliki range yang sempit dan berada pada nilai negative yaitu mulai dari -0.02 sampai 0.02 dengan keseluruhan nilai responden masuk ke dalam range sedang. Metode fuzzy tidak dapat mencapai nilai yang lebih tinggi atau nilai yang terdapat pada range “Tinggi” karena nilai awal selalu bernilai negatif walaupun responden memberikan jawaban yang benar. Selain itu loncatan nilai θ akhir pada metode fuzzy sangat kecil, sehingga tidak dapat mencapai nilai yang berada pada range “Tinggi” ketika *stopping rule* telah tercapai. Sedangkan loncatan nilai θ akhir metode IRT sangat bervariasi sesuai dengan tingkat kesulitan yang diberikan oleh sistem sehingga metode IRT dapat mencapai nilai maksimum sebelum *stopping rule* tercapai, sehingga bagi responden yang memiliki kemampuan tinggi pun dapat terukur dengan baik.

4. KESIMPULAN

Berdasarkan pengujian yang sudah dilakukan, baik menggunakan metode IRT ataupun metode Fuzzy sama-sama memberikan nilai yang sesuai dengan pengujian secara manual. Artinya metode IRT dan Fuzzy sudah dapat menggantikan proses pengujian secara manual dengan lebih efektif dan efisien. Dikatakan demikian karena kedua metode tersebut dapat memberikan hasil yang sesuai dengan hanya menyajikan butir soal sebanyak 37,5% dari soal yang disajikan secara manual.

Kesesuaian yang dicapai pada perbandingan pengujian manual dengan metode IRT sebesar 70%, sedangkan yang dicapai pada perbandingan manual dengan metode Fuzzy sebesar 80%. Namun range yang dihasilkan oleh metode IRT lebih luas dan bervariasi dibandingkan dengan metode fuzzy. Selain itu loncatan nilai θ akhir pada metode fuzzy sangat kecil, sehingga sangat sulit mencapai nilai yang berada pada range “Tinggi” ketika *stopping rule* telah tercapai. Sedangkan loncatan nilai θ akhir metode IRT sangat bervariasi sesuai dengan tingkat kesulitan yang diberikan oleh sistem sehingga metode IRT dapat mencapai nilai maksimum sebelum *stopping rule* tercapai.

Jadi berdasarkan poin-poin yang sudah dijelaskan di atas dapat dipastikan bahwa metode IRT lebih baik daripada metode Fuzzy untuk diterapkan pada sistem CAT.

REFERENSI

- [1] S. Hadi, “Pengembangan *Computerized Adaptive Test* Berbasis Web,” Yogyakarta: Aswaja Pressindo, 2013.
- [2] A. Fatkhudin, M.F. Hidayatullah, “Membandingkan IRT Model 3PL Dengan IRT Model 4PL Untuk Penilaian Menggunakan *Computerized Adaptive Test*,” SAINTEKBU: Jurnal Sains dan Teknologi, Volume 9 No. 1 November 2016.
- [3] L. M. Anggrenita, I. Muslim, M.R.A. Saf, “Pengembangan *Computerized Adaptive Test* (CAT) Menggunakan Algoritma Fuzzy Dengan Metode Tsukamoto Berbasis Web”, Jurnal Aksara Komputer Terapan Politeknik Caltex Riau, Vol. 5, No. 1, Tahun 2016.
- [4] L. Erawan, S. Santosa, “*Computer-Adaptive Test* Dengan Pendekatan Item Response Theory Satu Parameter,” Jurnal Teknologi Informasi, Volume 5 Nomor 2, Oktober 2009, ISSN 1414-9999.

- [5] *National Council on Measurement in Education.(n.d.). Glossary of Important Assessment and Measurement Terms.* Retrieved October 27, 2017 from http://www.ncme.org/ncme/NCME/Resource_Center/Glossary/NCME/Resource_Center/Glossary1.aspx?hkey=4bb87415-44dc-4088-9ed9-e8515326a061#anchorA
- [6] Winarno, 2012, *Pengembangan Computerized Adaptive Testing (CAT) Menggunakan Metode Pohon Segitiga Keputusan*, Jurnal Penelitian dan Evaluasi Pendidikan, Tahun 16, Nomor 2, 2012.
- [7] Hidayat, A. N., Nuril, A. T., Anindita, A., 2017, *Pengembangan Web Tool Computerized Adaptive Test Dengan Algoritma K-Means Untuk Clustering Soal Dan Pemilihan Butir Test Dengan Fuzzy Logic*, Seminar Nasional Teknologi Informasi dan Multimedia 2017, STMIK AMIKOM Yogyakarta, 4 Februari 2017, ISSN: 2302-3805.