

Implementasi *Item Response Theory Model Three-Parameter Logistics* Pada Aplikasi *Computerized Adaptive Test*

Muhammad Ruslan Maulani¹, Supriady²

Teknik Informatika, Politeknik Pos
Jl. Sariasih No. 54, Bandung, Indonesia
Email : ruslanmaulani@poltekpos.ac.id¹, supriady@poltekpos.ac.id²

Abstract

Paper and Pencils Tests (PPTs) are exams that provide test takers with questions of the same level of difficulty. As a result, the test becomes longer, and the questions given are not informative so that the examinees find it difficult or find it too easy to do it. The purpose of this study is to create a computer-based test system application that can measure the ability of test takers, where the level of difficulty of the items is adjusted to the ability level of the test takers. The research method carried out in this study consists of several stages, such as: the preparation, the analysis and design, the implementation and testing, and the reporting. While the method used to estimate the items using the Maximum Likelihood Estimation (MLE) method and the method used for the Computerized Adaptive Test (CAT) system, namely the Item Response Theory (IRT) method, the three-parameter logistic (3PL) logistic model. The result of this research is a web-based CAT system application that can be used as an online exam system and a measuring tool to determine the ability of the examinees.

Keywords: computerized adaptive test. item response theory, maximum likelihood estimation, paper and pencils test.

Abstrak

*Paper and Pencils Test (PPTs) merupakan ujian yang memberikan pertanyaan dengan tingkat kesulitan yang sama kepada peserta ujian. Akibatnya, tes menjadi lebih lama dan pertanyaan yang diberikan tidak informatif sehingga peserta ujian merasa kesulitan atau merasa terlalu mudah dalam mengerjakannya. Tujuan dari penelitian ini yaitu membuat aplikasi sistem ujian berbasis komputer yang dapat mengukur kemampuan peserta tes, dimana tingkat kesukaran butir soal disesuaikan dengan tingkat kemampuan dari peserta tes. Metode penelitian yang dilakukan dalam penelitian ini terdiri dari beberapa tahap, yaitu: tahap persiapan, tahap analisis dan perancangan, tahap implementasi dan pengujian, serta tahap pelaporan. Sedangkan metode yang digunakan untuk melakukan proses estimasi terhadap butir soal dengan menggunakan metode *Maximum Likelihood Estimation* (MLE) dan metode yang digunakan untuk sistem *Computerized Adaptive Test* (CAT) yaitu metode *Item Response Theory* (IRT) model logistik *three parameter logistic* (3PL). Hasil dari penelitian ini yaitu aplikasi sistem CAT berbasis web yang dapat dijadikan sebagai sistem ujian online dan alat ukur untuk menentukan kemampuan (*ability*) peserta ujian.*

Kata kunci: computerized adaptive test. item response theory, maximum likelihood estimation, paper and pencils test.

1. Pendahuluan

Penerapan teknologi informasi telah mengalami banyak perubahan terutama dalam dunia pendidikan. Salah satu dampak dari perkembangan teknologi yaitu komputer telah berevolusi menjadi alat yang dapat meningkatkan akurasi, efisiensi, antarmuka dan mekanisme umpan balik dari ujian [1]. Dengan adanya ujian seorang pengajar akan mengetahui apakah seorang siswa telah mencapai level yang tepat sesuai dengan pengetahuannya [2]. Evaluasi merupakan salah satu komponen penting pembelajaran yang harus dilaksanakan dengan baik karena fungsinya dalam memantau jalannya proses pembelajaran [3]. Sebenarnya, mempelajari suatu proses penilaian tidak hanya akan menambah efektifitas pembelajaran terhadap kemampuan siswa, tetapi akan memberikan umpan balik kepada instruktur, sehingga dapat meningkatkan produk dan layanan pembelajaran, serta menentukan strategi dalam proses pembelajaran [4], [5] dan [6].

Komputer merupakan hasil dan kemajuan yang signifikan dalam pengelolaan ujian, mulai dari ujian tradisional sampai munculnya *Computerize Adaptive Testing* (CAT) [7]. Ujian tradisional merupakan ujian yang dilakukan dengan menggunakan kertas dan pencil atau dikenal dengan istilah *Paper and Pencils Test* (PPTs). PPTs merupakan ujian yang memberikan pertanyaan dengan tingkat kesulitan yang

sama kepada peserta ujian. Akibatnya, tes menjadi lebih lama dan pertanyaan yang diberikan tidak informatif sehingga peserta ujian merasa kesulitan atau merasa terlalu mudah dalam mengerjakannya [8]. CAT merupakan kombinasi dari teori pengukuran dan teknologi komputer untuk menyediakan pengukuran yang tepat dalam pembelajaran berbasis komputer [9].

Dalam proses pengukuran, pengukuran dilakukan untuk mengetahui kualitas pertanyaan dan kemampuan peserta ujian. Kualitas pertanyaan dan kemampuan peserta ujian akan terlihat dengan mengetahui jawaban setiap peserta terhadap soal ujian. Untuk mengukur kualitas pertanyaan pada setiap butir soal dapat diketahui melalui parameter tingkat kesulitan (*difficulty*), fungsi pembeda soal (*discrimination*) dan menjawab dengan ditebak (*guessing*) [10].

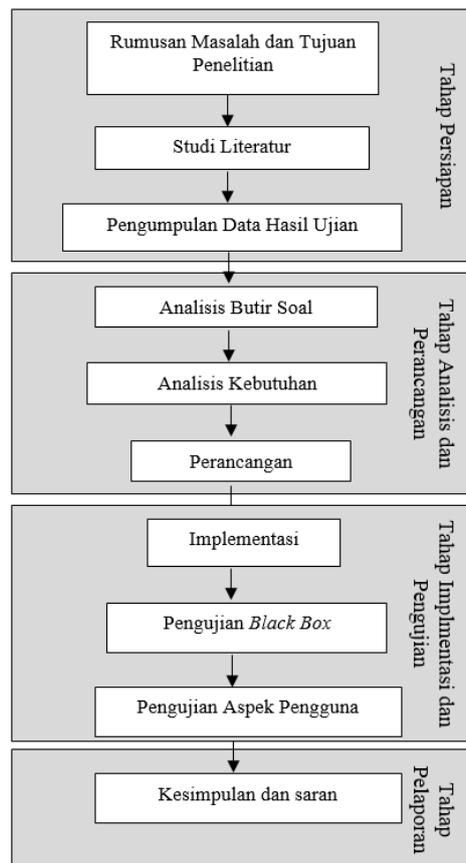
Teori yang mendasari CAT adalah *Item Response Theory* (IRT). CAT adalah ujian yang dikelola oleh komputer, dimana pemilihan pertanyaan pada nomor berikutnya dan keputusan untuk menghentikan ujian dilakukan secara dinamis berdasarkan kemampuan peserta yang dibuat dan dihitung selama ada interaksi dengan sistem [11]. Dengan cara ini, ketika pertanyaan yang diajukan ternyata terlalu sulit, maka pertanyaan berikutnya adalah pertanyaan dengan kesukaran yang lebih rendah, dan begitu juga sebaliknya [12].

Hal yang paling dasar mengenai CAT adalah memungkinkan tes disesuaikan dengan peserta ujian, yang memiliki keuntungan jelas yaitu secara signifikan dapat mengurangi waktu pengujian tanpa kehilangan akurasi [13], yang berarti evaluasi dapat dilakukan dengan cepat dan akurat. CAT memungkinkan item untuk dipilih berdasarkan tanggapan peserta terhadap item sebelumnya, memodifikasi tes untuk peserta tes [11]. Karena keunggulan ini, pengujian CAT telah berkembang pesat dalam beberapa dekade terakhir [14], terutama di berbagai bidang evaluasi seperti kepribadian wirausaha pada orang muda [11], kepribadian dari model Lima Besar [15] dan iklim organisasi [16]. Evaluasi hasil pembelajaran mungkin memerlukan opsi jawaban yang lebih kompleks dan terbuka [2]. Beberapa sistem, seperti Concerto [17] atau SIETTE [2], memungkinkan jenis item dengan menyediakan butir soal dalam bentuk HTML yang sederhana. Namun, item yang memerlukan presentasi dan evaluasi jawaban yang terdiri dari beberapa komponen (misalnya beberapa bidang input), menemukan beberapa kesalahan dalam gambar secara konseptual. IRT dapat diterapkan dalam ujian online dengan penilaian adaptif, materi yang diujikan yaitu mengenai pengenalan OOP yang tersedia dalam repository online [18].

Berdasarkan pemaparan di atas, maka penelitian ini penting dilakukan untuk mengetahui parameter butir soal dengan menerapkan *Item Response Theory*, menganalisis dan membuat perancangan prototype serta sistem CAT. Oleh karena itu, dalam penelitian ini peneliti mengambil judul “Implementasi *Item Response Theory Model Three-Parameter Logistics* Pada Aplikasi *Computerized Adaptive Test*”.

2. Metodologi

Metode penelitian yang dilakukan dalam penelitian ini terdiri dari beberapa tahap, yaitu: tahap persiapan, tahap analisis dan perancangan, tahap implementasi dan pengujian, serta tahap pelaporan. Metode penelitian yang dilakukan secara detail dapat dilihat pada Gambar 1.



Gambar 1. Metodologi Penelitian

2.1. Tahap Persiapan

Tahap persiapan merupakan tahap awal yang dilakukan oleh peneliti dalam kegiatan penelitian. Dalam tahap ini, peneliti melakukan identifikasi masalah dan tujuan dari penelitian yang akan dilakukan. Untuk mengidentifikasi masalah dan tujuan tersebut, peneliti melakukan pengumpulan data dan observasi pada objek atau tempat penelitian yang telah ditentukan. Selain itu, peneliti juga mencari studi literatur terhadap masalah yang telah ditentukan, untuk dijadikan sebagai sumber referensi dalam penelitian. Kegiatan tersebut dilakukan dengan tujuan agar peneliti dapat mengetahui suatu masalah, gagasan serta solusi dalam penelitian.

2.2. Tahap Analisis dan Perancangan

Pada tahap ini, peneliti melakukan proses analisis butir soal dan analisis terhadap kebutuhan. Kegiatan yang dilakukan yaitu:

- Analisis butir soal, melakukan proses estimasi parameter terhadap butir soal dengan menggunakan metode *Maximum Likelihood Estimation*. Dalam proses ini peneliti menggunakan software *expert system* yaitu XCalibre42.
- Analisis kebutuhan fungsional, mengenai analisis terhadap fungsi aplikasi dan batasan – batasan perangkat yang digunakan serta analisis kebutuhan *hardware* dan *software*.

Setelah melakukan tahap analisis, kemudian dilakukan perancangan sistem. Perancangan sistem dilakukan dengan mengacu pada standar *Rational Unified Process* [19] yaitu menggunakan pemodelan *Unified Modeling Language* (UML) [20] dan mengacu pada standar ISO 9241-11 [21]. Kegiatan yang dilakukan dalam tahap ini yaitu: merancang arsitektur sistem, merancang struktur kelas, merancang sub sistem dan komponen yang akan digunakan serta merancang tampilan antarmuka aplikasi.

2.3. Tahap Implementasi dan Pengujian

Tahap implementasi merupakan penerapan hasil rancangan dengan menggunakan bahasa pemrograman berbasis web. Dalam penelitian ini, peneliti menggunakan Bahasa pemrograman PHP-MySQL yang diintegrasikan dengan *Learning Management System Moodle*. Sedangkan pada tahap pengujian akan dilakukan pengujian fungsional, meliputi pengujian fungsi, fitur dan penanganan kesalahan. Pengujian ini menggunakan pengujian *black box*.

2.4. Tahap Pelaporan

Tahap ini merupakan tahap akhir yang dilakukan oleh peneliti. Kegiatan yang dilakukan oleh peneliti yaitu membuat laporan penelitian, membuat kesimpulan dari hasil pelaksanaan penelitian dan saran untuk penelitian berikutnya.

3. Hasil dan Pembahasan

3.1. Analisis Butir Soal

Analisis terhadap butir soal dilakukan hanya pada item butir soal CISCO chapter 2, 3 dan 4. Pada tiap chapter, data respons dikalibrasi dengan metode *marginal maximum likelihood (MML)* dengan menerapkan logistik 3-parameter (3PL). Adapun hasilnya dapat dilihat pada tabel 1, tabel 2 dan tabel 3.

Tabel 1. Parameter butir soal pada chapter 2

Item ID	a	b	c
1	1.315	0.2557	0.1781
2	1.6409	0.0344	0.1765
3	1.4639	0.2606	0.3066
4	1.7113	-0.098	0.242
5	1.3288	0.1681	0.2462
6	1.4701	0.0948	0.2438
7	1.5779	0.4775	0.2443
8	1.6684	0.1032	0.2403
9	1.5163	0.0692	0.243
10	1.3087	0.2573	0.2424
11	1.3549	0.3058	0.2497
12	1.5534	-0.287	0.2443
13	1.3954	-0.4123	0.2482
14	1.3916	0.041	0.2458
15	1.2971	-0.9582	0.2505
16	1.4509	0.6223	0.2461
17	1.4487	0.0924	0.2468
18	1.5436	-0.2854	0.2443
19	1.3368	-0.4937	0.2498
20	1.4615	0.6489	0.25
21	1.4471	-0.4398	0.2475

Keterangan: a = daya beda (*discrimination*); b = kesukaran (*difficulty*); dan c = tebakan (*guessing*)

Hasil kalibrasi item butir soal chapter 2 dalam Tabel 3.1 menampakkan bahwa parameter daya beda (a) tertinggi dimiliki oleh item nomor 4 dengan $a_4 = 1.7113$, sementara terendah dimiliki oleh item nomor 15

dengan $a_{15} = 1.2971$. Tingkat kesukaran tertinggi dimiliki oleh item nomor 20 dengan $b_{20} = 0.6489$ dan tingkat kesukaran terendah dimiliki oleh item nomor 15 dengan $b_{15} = -0.9582$. Sedangkan pada parameter tebakan, tertinggi pada item nomor 3 dengan $c_3 = 0.3066$ dan terendah pada nomor 2 dengan $c_2 = 0.1765$.

Tabel 2. *Parameter butir soal pada chapter 3*

Item ID	a	b	c
1	0.6827	-1.1353	0.2121
2	1.003	-0.7974	0.2086
3	1.0024	-0.8662	0.336
4	1.0863	-0.9968	0.2456
5	0.942	-1.1354	0.2477
6	1.0326	-1.0464	0.2482
7	0.9478	-0.5819	0.2483
8	0.9544	-1.2362	0.2491
9	0.9937	-1.3443	0.2503
10	1.0482	-1.7903	0.2477
11	0.8436	-0.7988	0.2483
12	0.9852	-0.8886	0.2481
13	0.8613	-0.629	0.2483
14	0.9218	-1.0667	0.2488
15	0.9579	-2.2281	0.2501
16	0.7691	-0.744	0.2504
17	1.076	-1.1618	0.2459
18	0.8193	-1.6419	0.251
19	0.9507	-1.1817	0.2466
20	0.8223	-0.7221	0.2469

Keterangan: a = daya beda (*discrimination*); b = kesukaran (*difficulty*); dan c = tebakan (*guessing*)

Hasil kalibrasi item butir soal chapter 3 dalam Tabel 3.2 menampakkan bahwa parameter daya beda (a) tertinggi dimiliki oleh item nomor 4 dengan $a_4 = 1.0863$, sementara terendah dimiliki oleh item nomor 1 dengan $a_1 = 0.6827$. Tingkat kesukaran tertinggi dimiliki oleh item nomor 7 dengan $b_7 = -0.5819$ dan tingkat kesukaran terendah dimiliki oleh item nomor 10 dengan $b_{10} = -1.7903$. Sedangkan pada parameter tebakan, tertinggi pada item nomor 3 dengan $c_3 = 0.3360$ dan terendah pada nomor 2 dengan $c_2 = 0.2086$.

Tabel 3. *Parameter butir soal pada chapter 4*

Item ID	a	b	c
1	0.2891	0.5134	0.2809
2	0.5215	-0.9788	0.2815
3	0.9005	-0.6743	0.3799
4	1.6006	-0.8505	0.2455
5	1.451	-0.8806	0.2484
6	1.7448	-0.6103	0.2489
7	1.5124	-0.7286	0.2492
8	1.4532	-0.7155	0.2499
9	1.2633	-0.839	0.2505
10	1.5822	-0.7399	0.2486

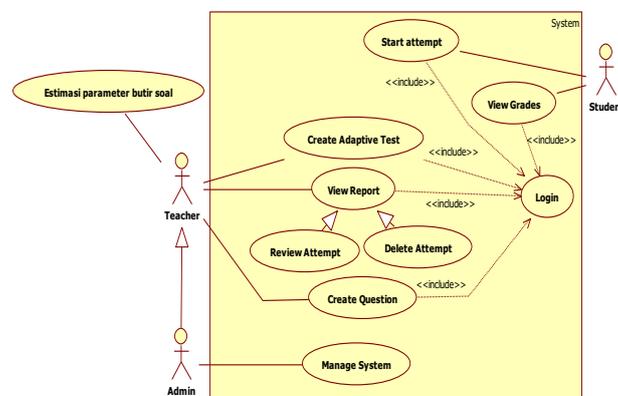
11	1.686	-0.5274	0.2454
12	1.4814	-1.3138	0.25
13	1.4866	-0.9594	0.249
14	1.5192	-0.3634	0.2455
15	1.1241	-0.8836	0.2534
16	1.5686	-1.3496	0.2488
17	1.6903	-0.4443	0.2451
18	1.1464	-0.9091	0.2524
19	1.7106	-0.7115	0.2458
20	1.266	-0.2458	0.2479

Keterangan: a = daya beda (*discrimination*); b = kesukaran (*difficulty*); dan c = tebakan (*guessing*)

Hasil kalibrasi item butir soal chapter 4 dalam Tabel 3.3 menampakkan bahwa parameter daya beda (a) tertinggi dimiliki oleh item nomor 6 dengan $a_6 = 1.7448$, sementara terendah dimiliki oleh item nomor 1 dengan $a_1 = 0.2891$. Tingkat kesukaran tertinggi dimiliki oleh item nomor 1 dengan $b_1 = 0.5134$ dan tingkat kesukaran terendah dimiliki oleh item nomor 16 dengan $b_{16} = -1.3496$. Sedangkan pada parameter tebakan, tertinggi pada item nomor 3 dengan $c_3 = 0.3799$ dan terendah pada nomor 11 dengan $c_{11} = 0.2454$.

3.2. Use Case Diagram

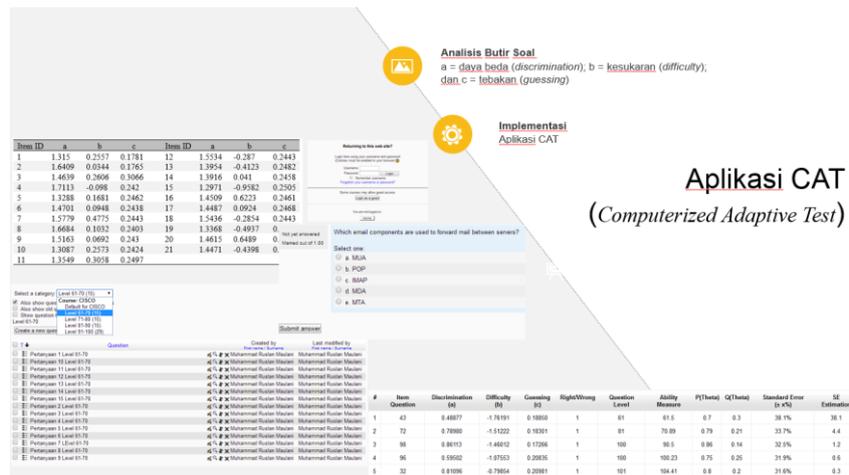
Use case diagram merupakan bagian utama dari pemodelan berorientasi objek yang menggunakan UML. Use case diagram menggambarkan fungsionalitas dari sebuah sistem sebagai sekumpulan skenario untuk menyelesaikan masalah tertentu dan mempresentasikan sebuah interaksi antara aktor dengan sistem. Use Case Diagram dari aplikasi CAT dapat dilihat pada gambar 2.



Gambar 2. Use Case Diagram Aplikasi CAT

3.3. Implementasi

Setelah dilakukan analisis dan perancangan, tahap berikutnya adalah melakukan proses implementasi sistem. Langkah implementasi merupakan langkah yang dilakukan sebagai integrasi dari perancangan sistem yang akan diterapkan pada aplikasi dan pengolahan data.



Gambar 3. Aplikasi CAT Model Three-Parameter Logistics

Halaman view report merupakan halaman yang berfungsi untuk menampilkan hasil ujian yang dilakukan oleh peserta. Halaman view report dapat diakses oleh user yang mempunyai hak akses sebagai teacher / admin. Pada halaman ini teacher dapat melihat kemampuan peserta (*ability*), *standard error* serta *number of attempt*.

Halaman review attempt merupakan halaman yang berfungsi untuk menampilkan scoring table dari hasil ujian yang dilakukan oleh peserta. Halaman review attempt dapat diakses oleh user yang mempunyai hak akses sebagai teacher / admin.

3.4. Hasil Pengujian

3.4.1. Keunggulan Aplikasi

Beberapa keunggulan yang dimiliki aplikasi ini, antara lain:

1. Tingkat kemampuan peserta diukur berdasarkan jawaban peserta terhadap soal bukan berdasarkan "true score".
2. Setiap pertanyaan yang diajukan disesuaikan dengan tingkat ability masing – masing peserta ujian, sehingga boleh jadi soal ujian yang dikerjakan oleh peserta ujian akan berbeda. Hal ini dapat mengurangi terjadinya kecurangan.
3. Layar komputer hanya menampilkan satu pertanyaan dalam satu waktu. Sehingga peserta tidak diperkenankan untuk melewati satu pertanyaan pun dan tidak pula diperkenankan untuk kembali ke pertanyaan yang telah dijawab sebelumnya. Hal ini bertujuan agar sistem dapat melakukan proses untuk memilih pertanyaan selanjutnya.

Pemilihan pertanyaan dan keputusan untuk menghentikan ujian ditentukan berdasarkan jawaban yang diberikan oleh masing – masing peserta ujian. Sehingga bisa jadi setiap peserta mengerjakan ujian dengan jumlah soal yang berbeda.

3.4.2. Kelemahan Aplikasi

Beberapa kelemahan yang dimiliki aplikasi ini, antara lain:

1. Kinerja aplikasi sangat bergantung pada konektivitas jaringan internet. Sehingga jika terjadi gangguan pada koneksi sistem tidak akan berjalan.
2. Aplikasi tidak dapat melakukan proses estimasi parameter butir soal. Dimana proses estimasi tersebut bertujuan untuk menghasilkan nilai parameter *discrimination*, *difficulty* dan *guessing*.

3.4.3. Analisis Manfaat

Aplikasi ini dapat membantu dalam dunia pendidikan terutama dalam proses pelaksanaan ujian, pengukuran kualitas soal serta kemampuan peserta ujian. Disamping itu aplikasi ini dapat ikut mendukung program pemerintah dalam dunia pendidikan yaitu dalam proses pelaksanaan ujian.

4. Kesimpulan

Peneliti melakukan proses estimasi terhadap butir soal pada matakuliah Administrasi Jaringan Komputer (CISCO) yaitu menggunakan metode *Maximum Likelihood Estimation* dengan bantuan software *expert system*, sehingga dapat diketahui parameter daya beda (*discrimination*), kesukaran (*difficulty*) dan tebakan (*guessing*) pada setiap butir soal. Dimana tingkat kesukaran tersebut dibagi kedalam kategori mudah, sedang dan sukar. Parameter yang diperoleh dari proses estimasi tersebut kemudian diimplementasikan dalam CAT berbasis IRT model *Three Parameter Logistics* (3PL). Hasil dari penelitian ini yaitu aplikasi sistem CAT berbasis web yang dapat dijadikan sebagai sistem ujian online dan alat ukur untuk menentukan kemampuan (*ability*) peserta ujian.

Daftar Pustaka

- [1] C. Wang and H. Lu, "Mediating effects of individuals' ability levels on the relationship of Reflective-Impulsive cognitive style and item response time in CAT," *Educ. Technol. Soc.*, vol. 21, no. 4, 2018.
- [2] R. Conejo, E. Guzmán, and M. Trella, "The SIETTE Automatic Assessment Environment," *Int. J. Artif. Intell. Educ.*, vol. 26, no. 1, 2016, doi: 10.1007/s40593-015-0078-4.
- [3] Norlaila, "Efektivitas Evaluasi Pembelajaran di Sekolah Dasar Islam Terpadu (SDIT) Ukhuwah Kota Banjarmasin," *Tashwir*, vol. 3, no. 5, 2015.
- [4] G. Lotito and G. Pirlo, "Item response theory for optimal questionnaire design," *J. E-Learning Knowl. Soc.*, vol. 9, no. 3, 2013, doi: 10.20368/1971-8829/820.
- [5] C. H. Lan, S. Graf, K. R. Lai, and K. Kinshuk, "Enrichment of peer assessment with agent negotiation," *IEEE Trans. Learn. Technol.*, vol. 4, no. 1, 2011, doi: 10.1109/TLT.2010.30.
- [6] C. Romero and S. Ventura, "Educational data mining and learning analytics: An updated survey," *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.*, vol. 10, no. 3, 2020, doi: 10.1002/widm.1355.
- [7] S. Oppl, F. Reisinger, A. Eckmaier, and C. Helm, "A flexible online platform for computerized adaptive testing," *Int. J. Educ. Technol. High. Educ.*, vol. 14, no. 1, 2017, doi: 10.1186/s41239-017-0039-0.
- [8] E. Istiyono, W. S. B. Dwandaru, R. Setiawan, and I. Megawati, "Developing of computerized adaptive testing to measure physics higher order thinking skills of senior high school students and its feasibility of use," *Eur. J. Educ. Res.*, vol. 9, no. 1, 2020, doi: 10.12973/eu-jer.9.1.91.
- [9] M. Rezaie and M. Golshan, "Computer Adaptive Test (CAT): Advantages and Limitations," *Int. J. Educ. Investig. Available online @ www.ijeionline.com*, vol. 2, no. 5, 2015.
- [10] S. Coronado, S. Sandoval-Bravo, P. L. Celso-Arellano, and A. Torres-Mata, "Competitive learning using a three-parameter logistic model," *Eur. J. Contemp. Educ.*, vol. 7, no. 3, 2018, doi: 10.13187/ejced.2018.3.448.
- [11] I. Pedrosa, J. Suárez-Álvarez, E. García-Cueto, and J. Muñiz, "A computerized adaptive test for enterprising personality assessment in youth," *Psicothema*, vol. 28, no. 4, 2016, doi: 10.7334/psicothema2016.68.
- [12] L. Peute, T. Scheeve, and M. Jaspers, "Classification and regression tree and computer adaptive testing in cardiac rehabilitation: Instrument validation study," *J. Med. Internet Res.*, vol. 22, no. 1, 2020, doi: 10.2196/12509.
- [13] Á. Postigo, M. Cuesta, I. Pedrosa, J. Muñiz, and E. García-Cueto, "Development of a computerized adaptive test to assess entrepreneurial personality," *Psicol. Reflex. e Crit.*, vol. 33, no. 1, 2020, doi: 10.1186/s41155-020-00144-x.
- [14] A. L. Zenisky and R. M. Luecht, "The future of computer-based testing: Some new paradigms.," in *Educational measurement: From foundations to future.*, 2016.
- [15] M. D. Nieto *et al.*, "Calibrating a new item pool to adaptively assess the Big Five," *Psicothema*, vol. 29, no. 3, 2017, doi: 10.7334/psicothema2016.391.
- [16] E. Peña-Suárez, F. Menéndez, F.-P. Eduardo, and J. Muñiz, "Computerized Adaptive Assessment of Organizational Climate," *An. Psicol.*, vol. 33, no. 1, 2016, doi: 10.6018/analesps.32.3.225921.
- [17] K. Scalise and D. D. Allen, "Use of open-source software for adaptive measurement: Concerto as an R-based computer adaptive development and delivery platform," *Br. J. Math. Stat. Psychol.*, vol. 68, no. 3, 2015, doi: 10.1111/bmsp.12057.
- [18] Y. L. P. Vega, J. C. G. Bolanos, G. M. F. Nieto, and S. M. Baldiris, "Application of item response theory (IRT) for the generation of adaptive assessments in an introductory course on object-oriented programming," 2012, doi: 10.1109/FIE.2012.6462377.

- [19] S. Shafiee, Y. Wautelet, L. Hvam, E. Sandrin, and C. Forza, "Scrum versus Rational Unified Process in facing the main challenges of product configuration systems development," *J. Syst. Softw.*, vol. 170, 2020, doi: 10.1016/j.jss.2020.110732.
- [20] U. M. Language, *Unified Modeling Language*, v2.5.1, no. Version 2.5.1. 2017.
- [21] K. Krisna, I. K. R. Arthana, and G. A. Pradnyana, "Pengujian Usability Pada Prototype Aplikasiwadaya Dengan Metode Usability Testing Mengadopsi Standar Iso 9241:11," *Ultim. J. Tek. Inform.*, vol. 11, no. 1, 2019, doi: 10.31937/ti.v11i1.1240.