

# Rancang Bangun Program Analisis Butir Soal Pilihan Ganda Sebagai Pendukung Proses Evaluasi Pembelajaran

Wisnu Wardoyo, Eko Suprpto

*Jurusan Teknik Elektro, Fakultas Teknik, Universitas Negeri Semarang, Indonesia*

*mas.wisnu99@gmail.com*

**Abstrak**— Salah satu bentuk evaluasi pendidikan adalah analisis butir soal. Akan tetapi proses analisis butir soal memiliki tingkat kesulitan tersendiri pada perhitungan matematis yang rumit khususnya untuk analisis butir soal pilihan ganda. Tujuan dari penelitian ini membangun sebuah perangkat lunak analisis butir soal pilihan ganda yang layak digunakan berdasarkan akurasi perhitungan analisis tingkat kesukaran, daya beda, keefektifan dan kualitas pengecoh, homogenitas, dan reliabilitas tes serta teruji sebagai perangkat lunak yang berkualitas. Perangkat lunak ini kemudian dibangun menggunakan bahasa pemrograman JAVA SDK 1.7. Metode penelitian yang digunakan adalah *Research and Development* melalui tahap analisis masalah, analisis kebutuhan, desain software, validasi desain, revisi desain, uji pakar, revisi 1, uji terbatas, revisi 2, uji diperluas dan revisi 3. Objek penelitian adalah perangkat lunak itu sendiri yang kemudian akan diujikan melalui uji kehandalan, uji Black-box, uji pakar, uji terbatas dan uji diperluas. Instrumen penelitian ini menggunakan angket. Hasil Penelitian menunjukkan bahwa perangkat lunak yang dikembangkan layak digunakan menganalisis butir soal pilihan ganda dan teruji sebagai sebuah perangkat lunak berdasarkan isi, tampilan, navigasi, keandalan dan kebermanfaatan perangkat lunak.

**Kata kunci**— evaluasi pendidikan, analisis butir soal, pilihan ganda, *Research and Development*.

## I. PENDAHULUAN

Salah satu cita – cita bangsa Indonesia yang tertuang dalam naskah pembukaan UUD 1945 adalah mencerdaskan kehidupan bangsa. Dalam rangka mencerdaskan kehidupan bangsa perlu ditempuh langkah – langkah konsisten melalui pendidikan. Pendidikan atau program pendidikan merupakan suatu aktivitas sadar atau sengaja yang diarahkan untuk mencapai cita – cita bangsa yaitu mencerdaskan bangsa. Untuk mengetahui keefektifan dan keefisienan sebuah program dalam mencapai tujuan maka dilakukanlah sebuah evaluasi terhadap komponen – komponen dan proses kerjanya sehingga apabila terdapat kegagalan dalam proses penyelenggaraan program tersebut dapat ditelusuri komponen dan proses mana yang menjadi sumber kegagalan.

Undang-undang Republik Indonesia nomor 20 tahun 2003, pada bab XVI pasal 57 sampai dengan 59 tentang Evaluasi menyatakan bahwa dalam rangka pengendalian mutu pendidikan secara nasional dilakukan evaluasi sebagai bentuk akuntabilitas penyelenggara pendidikan kepada pihak-pihak yang berkepentingan. Dengan kata lain evaluasi terhadap program pendidikan mutlak wajib dilakukan oleh setiap penyelenggara pendidikan khususnya guru. Guru dituntut kemampuannya melaksanakan evaluasi yang baik, dan secara otomatis dituntut membuat tes hasil belajar (THB) yang baik pula.

Salah satu bentuk evaluasi pendidikan adalah analisis butir soal. Akan tetapi, suatu laporan yang diturunkan oleh Dinas Pendidikan dan Kebudayaan Provinsi Jawa Tengah

memetakan kemampuan guru, menunjukkan bahwa kemampuan guru dalam menyusun butir soal kurang kompeten (Dinas Dikbud, 2011). Guru cenderung mengabaikan proses analisis butir yang mengedepankan prinsip-prinsip konstruksi tes. Akibatnya, kualitas tes menjadi tidak terstandar.

Berdasarkan studi yang dilakukan Eko Suprpto dkk, kebermutuan konstruksi butir soal semata-mata bukan terletak pada kemampuan guru menganalisis butir soal, akan tetapi terletak pada motivasi dan ketekunan guru yang harus menghitung secara manual aspek-aspek yang harus dikandung dalam suatu butir soal (Eko Suprpto dkk, 2013).

Berdasarkan studi tersebut maka perlu adanya seperangkat program yang bersifat melengkapi dan mendukung peningkatan kinerja guru. Penelitian ini menawarkan solusi untuk memberi alternatif bagi guru agar terbantu dalam melakukan analisis butir soal tes dengan mudah, efektif dan efisien melalui sebuah perangkat / program aplikasi komputer.

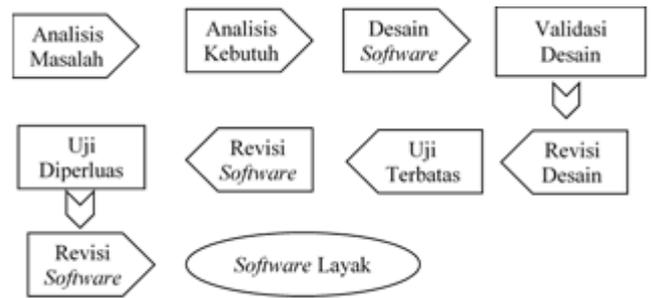
Fokus utama permasalahan yang akan diteliti pada penelitian ini adalah:

1. Apakah perangkat lunak analisis butir soal pilihan ganda yang dibangun mempunyai fungsionalitas yang baik dalam menganalisis butir soal pilihan ganda?
2. Apakah perangkat lunak analisis butir soal memiliki keunggulan dibanding perangkat lunak serupa yang terlebih dahulu sudah berada di pasaran?

Tujuan penelitian ini adalah sebagai berikut:

1. Merancang dan membangun perangkat lunak analisis butir soal pilihan ganda yang mempunyai fungsionalitas yang baik dalam menganalisis butir soal pilihan ganda.
2. Merancang dan membangun perangkat lunak analisis butir soal memiliki keunggulan dibanding perangkat lunak serupa yang terlebih dahulu sudah berada di pasaran.

Manfaat yang diharapkan dari penelitian ini dapat memberikan kemudahan bagi tenaga pengajar khususnya guru dalam melakukan proses evaluasi pembelajaran.



Gambar 1 Langkah-langkah penelitian

## II. METODE

Metode penelitian yang digunakan pada penelitian ini adalah metode *Research and Development* (R & D). Menurut Sugiyono (2009:297), metode penelitian dan pengembangan atau dalam bahasa Inggris disebut *Research and Development* adalah metode penelitian yang digunakan untuk menghasilkan produk tertentu, dan menguji keefektifan produk tersebut.

Produk tersebut tidak selalu berbentuk benda atau perangkat keras (*hardware*), seperti buku, modul, alat bantu pembelajaran di kelas atau di laboratorium, tetapi bisa juga software,

Dari analisis masalah dan analisis kebutuhan, proses analisis butir soal yang digunakan pada penelitian menggunakan teori klasik. Dalam teori klasik aspek-aspek yang dianalisis dalam suatu soal terdiri dari tingkat kesukaran, daya beda, keefektifan dan kualitas pengecoh, dan homogenitas, serta reliabilitas tes.

### 1. Tingkat Kesukaran

Tingkat kesukaran butir (*Item difficulty*) adalah angka yang menunjukkan besarnya proporsi peserta tes yang menjawab benar pada suatu butir (Mulyatiningsih, 2013: 172). Untuk mencari tingkat kesukaran soal menggunakan persamaan 1.

$$TK = \frac{nB}{N} \times 100\% \dots (1)$$

Keterangan:

nB = Jumlah peserta menjawab benar

N = Jumlah seluruh peserta

(Karno To, 2003:16).

Hasil penghitungan dari tingkat kesukaran selanjutnya diklasifikasikan menjadi beberapa kategori sesuai dengan hasil perhitungan, berikut adalah kriteria indeks tingkat kesukaran: atau program komputer yang digunakan untuk mempermudah pekerjaan tertentu.

Dalam pengembangan perangkat lunak ditempuh langkah-langkah penelitian sebagai Gambar 1.

TABEL I INDEKS TINGKAT KESUKARAN

Proporsi Benar	Kategori
0% – 15%	Sangat Sukar
16% – 30%	Sukar
31% – 70%	Sedang
71% – 85%	Mudah
86% – 100%	Sangat Mudah

### 2. Daya Beda

Daya pembeda butir soal berfungsi untuk melihat kemampuan butir soal dalam membedakan peserta tes antara peserta tes yang berkemampuan tinggi dengan peserta tes yang berkemampuan rendah (Mulyatiningsih, 2013:173). Untuk mencari daya beda soal menggunakan persamaan 2.

$$DP = \frac{BA - BB}{NA} \times 100\% \dots (2)$$

Keterangan:

BA = Jumlah Benar Kelompok A

BB = Jumlah Benar Kelompok B

NA = Jumlah Peserta A atau B

(Karno To, 2003:14).

Selanjutnya nilai daya beda akan dikategorikan berdasarkan pasaran nilainya. Berikut adalah indeks daya beda yang ditunjukkan pada Tabel II.

TABEL II INDEKS DAYA BEDA

Indeks Daya Pembeda	Kategori
Negatif – 9%	Sangat Buruk
10% - 19%	Buruk
20% - 29%	Cukup Baik
30% - 49%	Baik
50% ke atas	Sangat baik

### 3. Kualitas dan Keefektifan Pengecoh

Kualitas pengecoh dianggap baik bila jumlah siswa yang memilih pengecoh itu sama atau mendekati jumlah yang ideal (Karno To, 2003:17). Atau biasanya suatu soal dianggap pengecohnya efektif apabila masing-masing pengecoh dipilih minimal oleh satu peserta. Untuk mencari nilai kualitas pengecoh digunakan persamaan 3.

$$IPc = \frac{nPc}{(N - nB)/(Alt - 1)} \times 100\% \dots (3)$$

Keterangan :

- IPc = Indeks Pengecoh
- nPc = Jumlah siswa yang memilih pengecoh itu.
- N = Jumlah seluruh subjek yang ikut tes
- nB = Jumlah subjek yang menjawab benar pada butir tersebut.

(Karno To, 2003:17)

Berdasarkan nilainya indeks pengecoh dikategorikan dan ditunjukkan pada Tabel III.

TABEL III. INDEKS KUALITAS PENGECOH	
IPc	Kategori
76% - 125% (mendekati 100%)	Sangat Baik
51% - 75% atau 126% - 150%	Baik
26% - 50% atau 151% - 175%	Kurang Baik
0% - 25% atau 176% - 200%	Buruk
Lebih dari 200%	Sangat Buruk

#### 4. Homogenitas

Homogen tidaknya butir soal diketahui dengan menghitung koefisien korelasi antara skor tiap butir soal dengan skor total. Butir yang berkualitas adalah butir yang memiliki korelasi yang tinggi antara skor tiap butir dengan skor totalnya (Mulyatiningsih, 2013:175). Dalam mencari koefisien korelasi tersebut digunakan persamaan 4.

$$r_{xy} = \frac{n \Sigma XY - \Sigma X \cdot \Sigma Y}{\sqrt{[n \cdot \Sigma X^2 - (\Sigma X)^2][n \cdot \Sigma Y^2 - (\Sigma Y)^2]} \dots (4)$$

Keterangan :

- r = Koefisien Korelasi
- n = Jumlah peserta

(Karno To, 2003:10)

Butir soal yang homogen, koefisiennya korelasinya sama atau di atas batas signifikansi (batas kritis korelasi).

#### 5. Reliabilitas Tes

Reliabilitas tes menunjuk pada satu pengertian bahwa tes tersebut cukup dapat dipercaya sebagai tes yang baik dengan melihat tingkat konsistensi suatu tes. Dalam menguji reliabilitas tes pada perangkat lunak ini metode yang digunakan ada 2 yaitu, metode belah dua untuk jumlah soal genap metode KR21 untuk jumlah soal ganjil.

Untuk mencari koefisien reliabilitas menggunakan metode belah dua menggunakan persamaan 5.

$$r_{tt} = \frac{2 \times r_{xy}}{1 + r_{xy}} \dots (5)$$

Keterangan :

- rtt = Koefisien Reliabilitas Tes
- rx<sub>y</sub> = Koefisien korelasi ganjil – genap.

(Karno To, 2003:10)

Sedangkan untuk mencari koefisien reliabilitas pada KR21 dapat menggunakan persamaan 6.

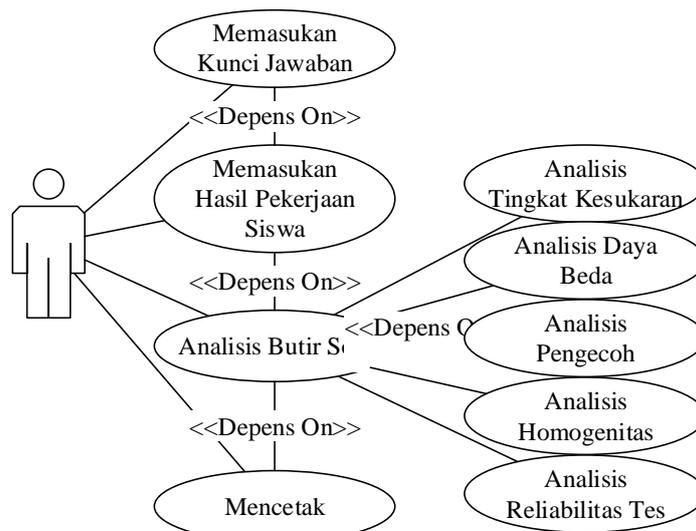
$$r_{11} = \left( \frac{n}{n-1} \right) \left\{ 1 - \frac{M(n-M)}{n s_1^2} \right\} \dots (6)$$

Keterangan :

- n = Jumlah butir
- M = Rata-rata skor total
- s<sub>1</sub><sup>2</sup> = Varian total

(Purwanto, 2010:171)

Tahapan selanjutnya merupakan desain software. Desain software diawali dengan analisis kebutuhan software. Analisis dilakukan untuk mengetahui kebutuhan software sehingga dapat memecahkan masalah yang ada. Dasar dari masalah tersebut adalah proses analisis butir soal yang dilakukan secara manual. Proses tersebut digambarkan dalam diagram aktivitas pada gambar 2.



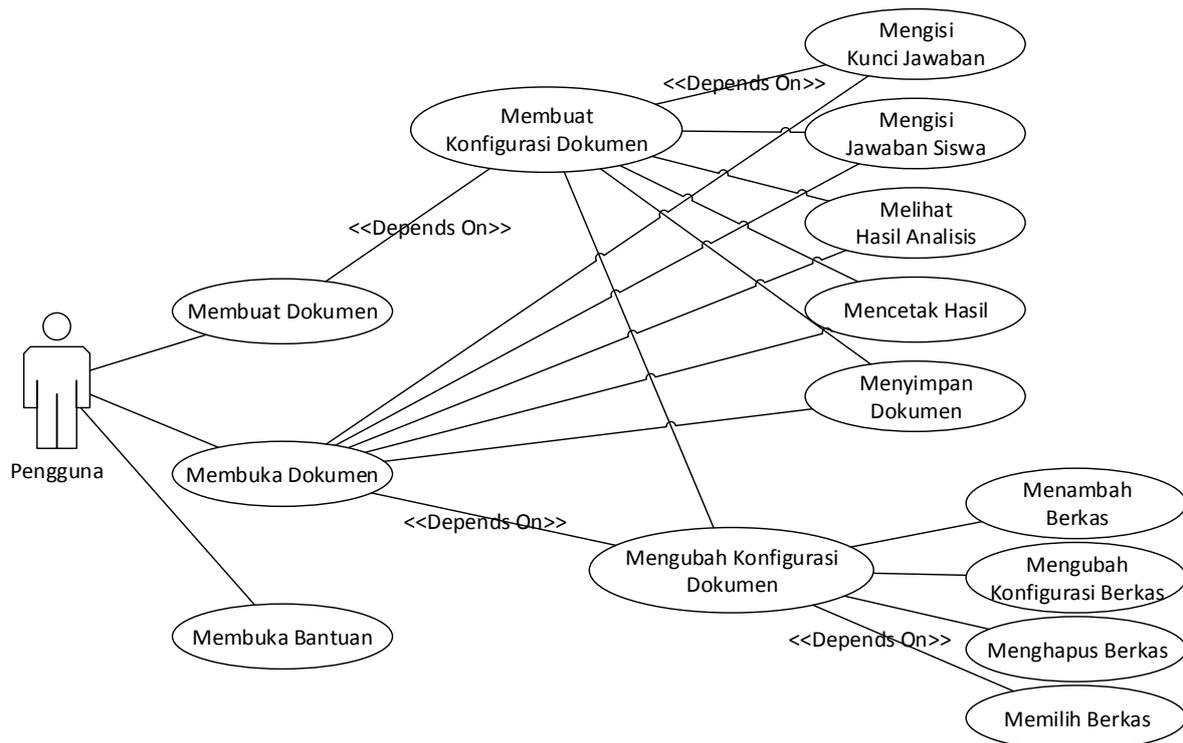
Gambar 2 Diagram Aktivitas Proses Analisis Butir Soal Secara Manual

Dari diagram aktivitas tersebut didapat permasalahan sebagai berikut:

1. Sistem analisis butir soal manual memakan waktu yang lama dalam proses penghitungannya.
2. Sistem analisis butir soal manual berpotensi terjadi kesalahan penghitungan yang dipengaruhi beberapa faktor dari luar.
3. Hasil dari sistem analisis soal manual selanjutnya tidak dapat dimanajemen dengan baik.
4. Memicu turunnya motivasi atau semangat guru untuk menganalisis butir soal karena kerumitannya.

Maka dirancanglah suatu perangkat lunak yang dapat permasalahan permasalahan tersebut. Perancangan perangkat lunak menggunakan perancangan model-driven dengan perancangan berbasis objek yang dibantu alat perancangan berupa UML.

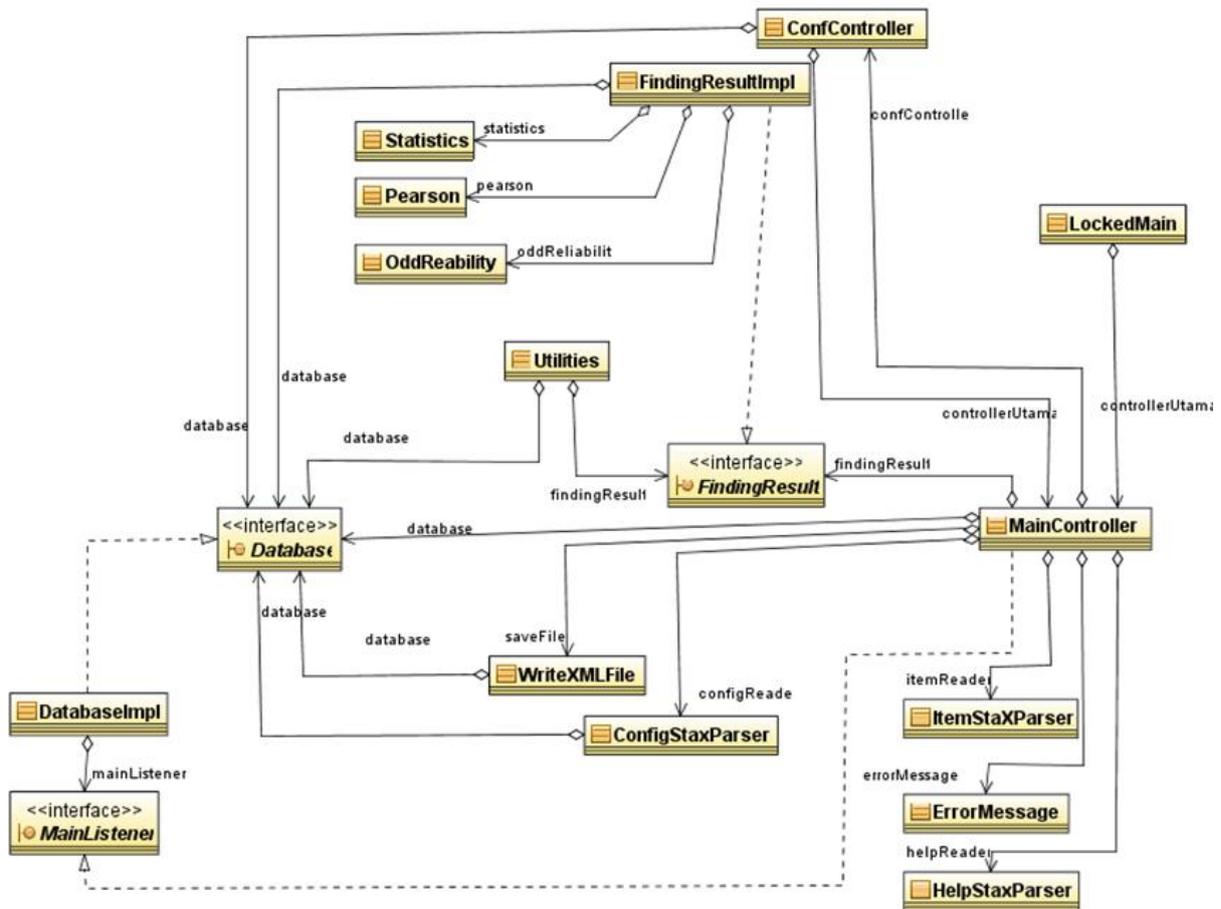
Perancangan model ini dahulu dengan perancangan use case diagram atau diagram yang menjelaskan tentang hubungan sistem dan fungsi yang ada dalam sistem tersebut. Berikut adalah rancangan perangkat lunak dalam bentuk use case diagram yang ditunjukkan pada Gambar 3.



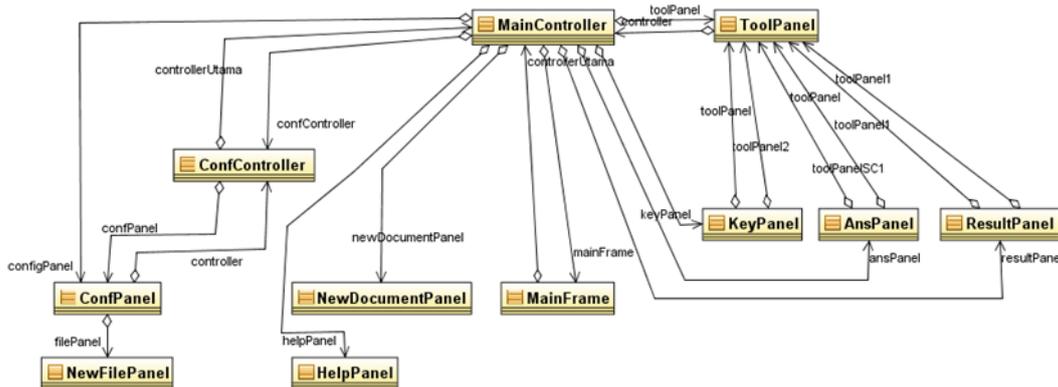
Gambar 3 Use Case Diagram Perancangan Perangkat Lunak

Dari masing-masing use case dalam use case diagram selanjutnya diimplementasikan menjadi kelas-kelas yang selanjutnya digunakan untuk dasar pembuatan perangkat lunak. Kelas-kelas tersebut saling berhubungan membentuk class diagram. Class diagram merepresentasikan objek-objek

dalam sistem saling berhubungan. Class diagram ini di rancangan model MVC (*model-view-controller*). Berikut rancangan perangkat lunak ditunjukkan pada Gambar 4 dan Gambar 5.



Gambar 4 Class Diagram Model-Controller



Gambar 5 Class Diagram View-Controller

Dari tahap tersebut dilanjutkan dengan tahap validasi desain. Dalam validasi desain terdapat tiga urutan yaitu uji kehandalan produk, uji blackbox dan uji pakar. Setelah validasi desain dilanjutkan dengan uji terbatas dan implementasi pengguna atau uji diperluas.

### III HASIL DAN PEMBAHASAN

Hasil dari penelitian ini merupakan produk yang dikembangkan dengan nama "AnisSo". Pada tahap validasi

desain, pengujian yang pertama adalah uji kehandalan produk. Pada uji kehandalan AnisSo telah menganalisis sebuah perangkat tes berupa soal ujian akhir semester(UAS) semester genap 2013/2014 mata pelajaran Matematika untuk kelas XI SMK. Dari hasil pengujian tersebut AnisSo mampu menunjukkan keakuratannya dalam menganalisis butir soal. Dari hasil uji tersebut didapatkan rekapitulasi yang ditunjukkan pada Tabel IV.

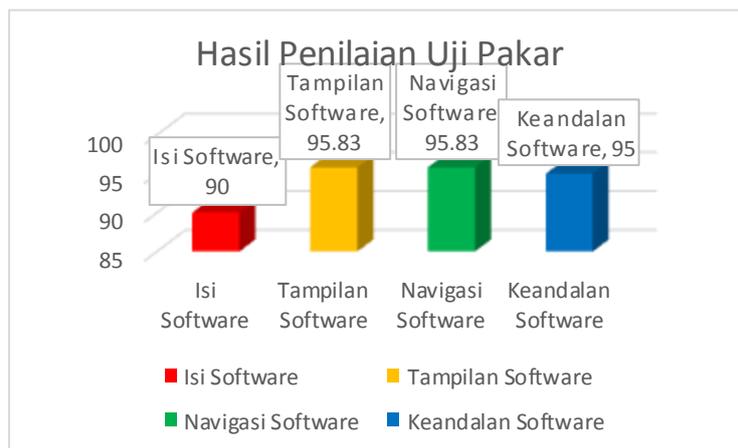
TABEL IV HASIL PENGUJIAN PERANGKAT TES DALAM UJI KEHANDALAN PRODUK

Aspek	Jumlah Soal	Keakuratan Penghitungan
<b>Tingkat kesukaran</b>		100%
Sangat Sukar	0	
Sukar	3	
Sedang	13	
Mudah	10	
Sangat Mudah	14	
<b>Daya Beda</b>		100%
Buruk	17	
Cukup	13	
Baik	8	
Baik Sekali	2	
<b>Efektivitas Pengecoh</b>		100%
Efektif	18	

Tidak Efektif	22	
<b>Homogenitas</b>		100%
Signifikan	9	
Sangat Signifikan	9	
Tidak Homogen	22	
<b>Reliabilitas Tes</b>		100%
Reliabilitas	0,69	

Dari hasil pengujian tersebut AnisSo juga menyarankan agar 21 soal dibuang, 11 soal direvisi dan 8 soal untuk dipertahankan.

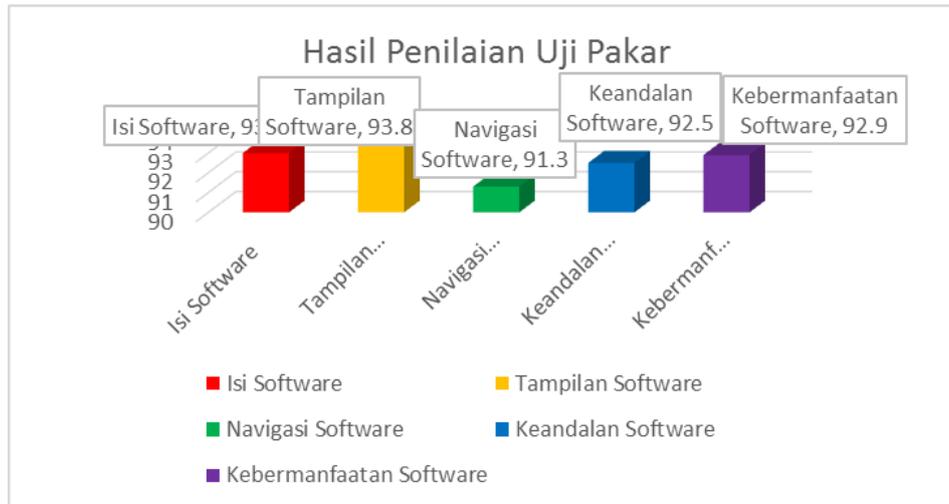
Pada tahap uji *blackbox* AnisSo mampu menjalankan semua fungsi yang telah dirancang dengan baik sesuai skenario yang telah dirancang. Pada tahapan ini AnisSo sudah siap untuk diujikan kepada pakar untuk mengetahui kualitas isi, tampilan, navigasi dan keandalan perangkat lunak. Uji pakar yang dilaksanakan di UNNES oleh tiga pakar menghasilkan persentase yang ditunjukkan pada Gambar 6.



Gambar 6 Hasil Penilaian Uji Pakar

Dari hasil uji pakar para pakar menilai keempat aspek perangkat lunak yaitu isi, tampilan, navigasi dan keandalan dalam kategori sangat baik. Hasil ini selanjutnya menjadi pijakan melanjutkan ke tahap uji terbatas. Uji terbatas dilaksanakan di SMK Negeri 7. Uji terbatas yang dimaksudkan melihat performa perangkat lunak dalam

menganalisis butir soal berdasarkan data nyata ini diikuti oleh 10 responden. Berikut adalah hasil uji terbatas terhadap kesepuluh responden tersebut yang ditunjukkan pada Gambar 7.



Gambar 7 Hasil penilaian Uji Terbatas

Hasil yang didapat hampir serupa dengan uji pakar, akan tetapi ada sedikit perbedaan antara uji pakar dan uji terbatas, pada uji terbatas terdapat satu aspek tambahan yang menjadi pertanyaan yaitu kebermanfaatan. Jika dilihat nilai dari aspek kebermanfaatan perangkat lunak yang telah dibangun memiliki kebermanfaatan yang sangat baik di mata para responden. Hal tersebut menjadi keunggulan tersendiri, dan menjadi sebuah modal awal untuk tahap uji selanjutnya.

Setelah dua uji yaitu validasi desain dan uji terbatas, AnisSo menerima berbagai perbaikan sehingga AnisSo memiliki banyak keunggulan. AnisSo memiliki beberapa keunggulan utama yang tidak dimiliki beberapa perangkat lunak serupa yang sudah ada sebelumnya. Keunggulan AnisSo adalah sebagai berikut:

1. Dapat dijalankan di banyak sistem operasi komputer seperti Windows, Mac OS dan Linux.
2. Mendukung sistem komputer dengan arsitektur 32bit dan 64bit.
3. Adanya fitur *copy-paste* yang mempermudah proses pemasukan data.

Selain dari 3 keunggulan utama pada AnisSo terdapat keunggulan lain pada AnisSo yaitu hasil cetak yang lebih baik, tampilan antar muka dengan ikon-ikon yang mudah dipahami, penerapan menu bar, dan penambahan fitur warning atau peringatan untuk beberapa kasus tertentu.

Jika dilihat dari hasil implementasi ke pengguna pada uji diperluas, semua responden sepakat bahwa AnisSo secara fungsionalitas sangat baik digunakan untuk menganalisis butir soal pilihan ganda. Dari aspek-aspek yang diujikan sebelumnya, mayoritas responden menyatakan tiap aspeknya dalam kategori sangat baik. Berikut adalah persentase tiap aspeknya menurut pengguna yang ditunjukkan pada Tabel V.

TABEL V TABEL PERSENTASE PENILAIAN PENGGUNA TERHADAP ASPEK PERANGKAT LUNAK

Aspek	Penilaian Pengguna
Isi Perangkat Lunak	76% Sangat Baik
	22% Baik
	2% Tidak Baik
Tampilan Perangkat Lunak	76% Sangat Baik
	24% Baik
Keandalan Perangkat Lunak	81% Sangat Baik
	19% Baik
Kebermanfaatan Perangkat Lunak	89% Sangat Baik
	11% Baik

Hal ini menjadi bukti bahwa teknologi komputer sangat bermanfaat bagi dunia pendidikan khususnya untuk mempermudah melakukan perhitungan yang rumit pada analisis butir soal. Dengan adanya pengembangan ini diharapkan menjadi suatu terobosan baru di dunia pendidikan, untuk dapat memotivasi guru melakukan analisis butir soal. Sehingga kompetensi guru semakin meningkat dan semakin memajukan pendidikan di Indonesia.

Walaupun demikian, penggunaan komputer sebagai alat bantu pendidikan di Indonesia belum begitu maksimal. Hal ini dikarenakan formasi guru di Indonesia didominasi oleh guru dengan usia yang relatif tua. Dalam beberapa kasus dalam penelitian dijumpai bahwa, penerimaan responden akan satu perangkat lunak baru butuh waktu dan latihan yang rutin serta motivasi perubahan dari diri responden. Kebanyakan yang sudah terbiasa tidak melakukan analisis tetap akan tidak melaksanakan analisis walaupun sudah ada alat bantu yang memadai. Hal tersebut karena rendahnya motivasi serta banyaknya tuntutan guru sehingga mereka tidak mempunyai cukup waktu untuk melakukan hal tersebut.

#### IV KESIMPULAN DAN SARAN

##### A. Kesimpulan

Berdasarkan hasil penelitian dan pembahasan, maka dapat diperoleh kesimpulan bahwa:

1. Perangkat lunak AnisSo baik secara fungsionalitas digunakan sebagai alat bantu analisis butir soal pilihan ganda.
2. Perangkat lunak AnisSo teruji memiliki akurasi yang tepat dalam menghitung aspek-aspek yang diisyaratkan prinsip analisis butir soal pilihan ganda.
3. Perangkat lunak program analisis butir soal memiliki aspek isi, tampilan, navigasi, kehandalan dan kebermanfaatan yang cenderung sangat baik.
4. Perangkat lunak program analisis butir soal pilihan ganda AnisSo memiliki keunggulan dalam segi arsitektur dan tampilan perangkat lunak dibanding perangkat lunak lainnya.

##### B. Saran

Berdasarkan kesimpulan-kesimpulan yang telah dikemukakan, dapat diajukan saran-saran pengembangan perangkat lunak lebih lanjut:

1. Diperlukan sosialisasi lebih lanjut tentang penggunaan perangkat lunak AnisSo. Harapannya dengan sosialisasi lebih lanjut, pengguna akan
2. mengetahui adanya perangkat lunak yang membantu menganalisis soal pilihan ganda dan mulai menggunakannya.
3. Pengembangan lebih lanjut terhadap perangkat lunak AnisSo. Sehingga perangkat lunak ini lebih sempurna baik secara prinsip analisis butir soal pilihan ganda atau fungsionalitas perangkat lunak.
4. Perlu adanya pengujian yang lebih komprehensif terhadap perangkat lunak AnisSo.

#### REFERENSI

- [1] Davis William. 1991. Sistem Pengelolaan Informasi. Jakarta: Erlangga.
- [2] Mulyatiningsih Endang. 2013. Metode Penelitian Terapan Bidang Pendidikan. Bandung : Alfabeta
- [3] Purnono Herry. 2005. Pengenalan Informatika Perspektif Teknik dan Lingkungan. Yogyakarta : Andi
- [4] Purwanto. 2011. Evaluasi Hasil Belajar. Yogyakarta: Pustaka Pelajar.
- [5] Sudjana. 11. Penilaian Hasil Proses Belajar Mengajar. Bandung : PT Remaja Rosdakarya
- [6] Sugiyono, 2009, Metode Penelitian Kuantitatif, Kualitatif dan R&D, Bandung : Alfabeta
- [7] Suprptono, Pribadi, Sugiarno. 2013. Artikel Penelitian Hibah Bersaing : Model Aplikasi Visual Basic Application for Excell (VBA) untuk Menganalisis Butir Tes pada Matapelajaran Rumpun Adaptip di SMK Negeri. Seramang.
- [8] *Undang-undang Republik Indonesia nomor 20 tahun 2003. Tentang Sistem Pendidikan Nasional.*
- [9] Badan penelitian dan Pengembangan Kementerian Pendidikan Nasional, 2010. Kajian Kemampuan Guru Mengelola Pembelajaran Berbasis Penelitian. Jakarta.