

Pengembangan Software Analisis Butir Klasik Untuk Calon Guru

Hari Purnomo Susanto¹⁾ , Fathurrahman²⁾ , Mega Isvandiana Purnamasari³⁾ , Khoirul Qudsiyah⁴⁾ , Johri Sabaryati⁵⁾ , Jero Budi Darmayasa⁶⁾ , Suciati⁷⁾ , Eko Wahyunanto Prihono⁸⁾ , Mobinta Kusuma⁹⁾

^{1,3,4)}STKIP PGRI Pacitan

²⁾Universitas Pendidikan Muhammadiyah Sorong

⁵⁾Universitas Muhammadiyah Mataram

^{6,7)}universitas Borneo Tarakan

⁸⁾Institut Agama Islam Negeri Ambon

⁹⁾Universitas Panca sakti Tegal

haripsusanto@stkippacitan.ac.id haripurnomosusanto@gmail.com

Abstrak: Sebagian besar perangkat lunak analisis butir berbasis *Classical Test Theory* (CTT) yang tersedia saat ini, baik berbayar maupun gratis, masih berorientasi pada pengguna ahli statistik dan tidak dilengkapi fitur interpretasi hasil analisis. Kondisi ini menyulitkan calon guru yang memiliki kemampuan beragam dalam memahami hasil evaluasi butir, karena mereka harus menafsirkan nilai-nilai statistik secara mandiri. Untuk menjawab kebutuhan tersebut, dikembangkan **software untuk analisis CTT**, sebuah perangkat lunak berbasis web berbahasa Indonesia yang dirancang khusus bagi calon guru agar dapat melakukan analisis butir secara praktis, edukatif, dan gratis. Pengembangan software pada penelitian ini mengadopsi model SDLC dengan lima langkah utama yaitu *planning*, *Analyst*, *Design*, *Implementation*, dan *system*. Software yang dihasilkan selanjutnya diberi nama Software Statistik a Sederhana seri *Classical Theory Test* (3SCTT) .Hasil pengembangan menunjukkan bahwa 3SCTT mampu menyajikan (1) hasil analisis butir secara lengkap;(2) interpretasi hasil analisis; (3) saran ; (4) software gratis dan (5) berbahasa indonesia. Selain berfungsi sebagai alat bantu analisis, 3SCTT juga berperan sebagai media pembelajaran CTT yang mendukung penguasaan konsep evaluasi pendidikan bagi calon guru. Dengan tampilan antarmuka yang sederhana dan fitur interpretasi otomatis, aplikasi ini membantu calon guru memahami makna setiap hasil analisis, bukan sekadar memperoleh angka. Ke depan, 3SCTT berpotensi dikembangkan lebih lanjut melalui integrasi dengan platform pembelajaran digital (LMS) dan penerapan kecerdasan buatan (AI) untuk rekomendasi revisi butir.

Kata Kunci : CTT, analisis butir, pengukuran, software, 3SCTT

Abstract:

Most existing software for item analysis based on the Classical Test Theory (CTT), whether paid or free, is still designed primarily for expert users in statistics and lacks features that provide interpretation of the analysis results. This limitation poses challenges for pre-service teachers with diverse abilities, as they must interpret statistical outputs independently. To address this issue, a CTT analysis software was developed a web-based application in Indonesian, specifically designed to enable pre-service teachers to conduct item analysis in a practical, educational, and free manner. The software development process adopted the System Development Life Cycle (SDLC) model, which consists of five main stages: Planning, Analysis, Design, Implementation, and System Testing. The resulting software was named

Software Statistika Sederhana seri Classical Test Theory (3SCTT). The development outcomes indicate that 3S-CTT is capable of providing: (1) comprehensive item analysis results; (2) interpretation of findings; (3) recommendations for improvement; (4) free accessibility; and (5) a fully Indonesian interface.

In addition to serving as an analytical tool, 3SCTT also functions as a learning medium for CTT, supporting pre-service teachers in mastering key concepts of educational assessment. With its simple user interface and automated interpretation features, the application helps users understand the meaning behind each statistical result, rather than merely obtaining numerical outputs. In the future, 3SCTT has the potential to be further developed through integration with Learning Management Systems (LMS) and the application of Artificial Intelligence (AI) to provide automatic item revision recommendations.

Keywords: *Classical Test Theory (CTT), item analysis, measurement, software, 3SCTT*

1. Pendahuluan

Dalam ranah teori pengukuran, *Classical Test Theory* (CTT) menempati posisi fundamental sebagai paradigma awal yang membentuk dasar pemahaman terhadap konsep pengukuran dalam pendidikan dan psikologi. Sebagai salah satu teori tertua dan paling berpengaruh, CTT telah memberikan kerangka konseptual yang kuat bagi lahirnya berbagai teori pengukuran modern yang lebih kompleks dan canggih (Hu et al., 2021; Sumaryanta, 2021). Walaupun tidak selalu diajarkan secara eksplisit melalui mata kuliah yang secara langsung berjudul “*Classical Test Theory*”, substansi dan prinsip-prinsip dasarnya tetap menjadi bagian integral dalam pembelajaran di tingkat sarjana, terutama pada bidang evaluasi pendidikan (Hayat, 2021). Keberlanjutan pengajaran teori ini mencerminkan signifikansinya dalam memberikan pemahaman mendasar mengenai sifat pengukuran, validitas, reliabilitas, serta kesalahan pengukuran yang menjadi fondasi bagi penerapan teori-teori pengukuran kontemporer.

Dalam praktik pengembangan instrumen pendidikan, CTT tetap menjadi pendekatan yang paling luas digunakan dibandingkan dengan teori-teori pengukuran modern lainnya seperti *Item Response Theory* (IRT) dan model Rasch. Dominasi penggunaan CTT tersebut tidak terlepas dari berbagai keunggulan yang dimilikinya, yang menjadikannya lebih praktis dan mudah diterapkan oleh pengembang tes (Sumaryanta, 2021). Beberapa keunggulan utama CTT meliputi kebutuhan jumlah sampel yang relatif kecil, prosedur analisis matematis yang sederhana, serta estimasi parameter yang mudah dipahami secara konseptual. Selain itu, penerapan CTT tidak menuntut analisis kompleks untuk menguji kesesuaian model terhadap data, sehingga menjadikannya pilihan yang efisien bagi peneliti maupun praktisi dalam konteks pendidikan.

Pada penerapan dalam pengembangan instrumen pendidikan, CTT masih menjadi teori yang paling sering digunakan dibandingkan teori pengukuran yang lain seperti Item Response theory (IRT) dan Rasch. Fakta ini disebabkan karena CTT memiliki kelebihan-kelebihan yang memudah pengembang tes dalam menggunakannya (Sumaryanta, 2021). Kelebihan-kelebihan tersebut diantaranya Sampel yang diperlukan cenderung kecil, perhitungan matematis yang digunakan sederhana, estimasi parameter secara konsep mudah dipahami, dan tidak memerlukan kajian mendalam untuk memastikan kecocokan model dengan data uji.

Beragam perangkat lunak telah dikembangkan untuk mendukung analisis *Classical Test Theory* (CTT). Beberapa di antaranya antara lain *ITEMAN*, *Quest*, program **R** dengan

paket *CTT* (R Core Team, 2023; Willse, 2018), *Jmetrik* (Meyer, 2014), serta *Klasika* (Hayat, 2021), dan masih terdapat sejumlah aplikasi lainnya yang memiliki fungsi serupa. Masing-masing perangkat lunak tersebut menawarkan keunggulan tersendiri dalam hal kemudahan penggunaan maupun keluaran analisis yang dihasilkan. Perangkat seperti *ITEMAN*, *Jmetrik*, dan *Klasika* dikenal memiliki antarmuka pengguna (*user interface*) yang intuitif sehingga memudahkan peneliti maupun praktisi dalam mengoperasikannya. Namun demikian, seluruh perangkat tersebut masih memiliki beberapa keterbatasan. Pertama, belum terdapat fitur yang secara otomatis membantu pengguna dalam menafsirkan hasil analisis *CTT*, sehingga potensi kesalahan interpretasi masih mungkin terjadi. Kedua, sebagian besar perangkat tersebut tidak menyediakan panduan penggunaan yang terintegrasi langsung di dalam aplikasi, sehingga pengguna perlu mempelajari petunjuk dari sumber eksternal. Ketiga, dari sisi bahasa, hanya *Klasika* yang menggunakan bahasa Indonesia, sementara aplikasi lainnya umumnya berbahasa Inggris.

Bagi calon guru, pendekatan *CTT* tetap menjadi pilihan utama dalam merancang dan menganalisis instrumen penilaian karena kemudahannya dalam aplikasi, seperti penghitungan reliabilitas dan indeks butir yang relatif sederhana. Namun, sejumlah kesulitan yang sering ditemukan beberapa kesulitan antara lain yaitu interpretasi hasil analisis, kesalahan entri data sehingga software menjadi error atau tidak mengeksekusi data. Tidak mengetahui apa yang harus dilakukan jika terdapat karakteristik butir yang kurang bagus atau distraktor yang buruk. Dalam hal ini calon guru tidak memiliki pengalaman apa yang harus diputuskan terhadap butir yang bermasalah.

Kesalahan dalam menafsirkan hasil analisis pengukuran merupakan hal yang krusial, karena dapat berimplikasi serius terhadap kesimpulan yang diambil serta keputusan yang dihasilkan. Dalam konteks pengukuran psikometrik, khususnya pada analisis instrumen atau butir tes, kesalahan interpretasi berpotensi menimbulkan pemahaman yang keliru terhadap konstruk, kemampuan, atau karakteristik yang hendak diukur. Kondisi tersebut dapat berdampak luas, mulai dari perancangan instrumen yang kurang tepat, pemilihan butir yang tidak representatif, hingga pengambilan keputusan pendidikan yang tidak adil terhadap individu yang diuji. Oleh karena itu, pemahaman konseptual yang kuat mengenai teori pengukuran serta keterampilan dalam menerapkan teknik analisis yang tepat menjadi sangat penting untuk menjamin hasil pengukuran yang valid dan dapat dipertanggungjawabkan secara ilmiah. Salah satu strategi untuk meminimalkan kesalahan interpretasi tersebut adalah dengan mengintegrasikan fitur panduan atau interpretasi hasil analisis secara otomatis pada perangkat lunak pengukuran, sebagaimana telah dikembangkan dalam *software 3SAR*(Susanto, Meifiani, et al., 2023) dan *irtawsi*(Susanto, Retnawati, et al., 2023).

Beragam perangkat lunak yang digunakan untuk analisis *Classical Test Theory* (*CTT*) menawarkan banyak fitur dan opsi yang terkadang justru menimbulkan kebingungan bagi pengguna, khususnya bagi mereka yang masih memiliki pengalaman terbatas dalam analisis psikometrik. Salah satu contoh permasalahan yang sering muncul adalah terjadinya kesalahan (*error*) pada saat penggunaan *ITEMAN*, yang umumnya disebabkan oleh pengaturan atau format data yang tidak sesuai dengan spesifikasi program. Kondisi ini menegaskan pentingnya pemahaman yang memadai mengenai prosedur kerja perangkat lunak yang digunakan, serta perlunya verifikasi hasil analisis secara cermat sebelum dijadikan dasar pengambilan keputusan atau kebijakan yang berkaitan dengan evaluasi pendidikan.

Selain pemahaman teknis, aspek bahasa yang digunakan oleh *software* juga memiliki peran penting dalam mendukung keakuratan penggunaan perangkat lunak analisis. Pemilihan bahasa yang sesuai dengan konteks dan tingkat literasi pengguna dapat mengurangi risiko

kebingungan maupun kesalahan interpretasi terhadap instruksi dan hasil keluaran. Penggunaan bahasa yang jelas, komunikatif, dan mudah dipahami akan membantu pengguna menavigasi berbagai fitur secara lebih efektif, meningkatkan rasa percaya diri, serta mempercepat proses pembelajaran bagi pengguna baru. Salah satu pendekatan yang dapat mengatasi kendala tersebut adalah penyediaan fitur pilihan bahasa pada perangkat lunak, sebagaimana diterapkan dalam *software irtawsi*, yang memungkinkan pengguna untuk beralih antara bahasa Inggris dan bahasa Indonesia (Susanto, Retnawati, et al., 2023).

Fakta-fakta yang telah diuraikan sebelumnya menunjukkan bahwa berbagai perangkat lunak analisis *Classical Test Theory* (CTT) masih menimbulkan sejumlah kendala yang dapat mengurangi kenyamanan dan efektivitas calon guru dalam melakukan analisis instrumen. Kondisi ini mengindikasikan perlunya pengembangan perangkat lunak analisis CTT yang mampu mengatasi keterbatasan yang terdapat pada aplikasi yang sudah ada. Inovasi dapat diwujudkan melalui penyediaan fitur-fitur yang mendukung kemudahan penggunaan, seperti penyajian hasil analisis yang disertai interpretasi otomatis, pilihan bahasa yang sesuai dengan konteks pengguna, serta panduan langkah-langkah penggunaan yang praktis dan intuitif.

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk mengembangkan perangkat lunak statistika sederhana berbasis *Classical Test Theory*. Perangkat lunak ini dirancang khusus untuk membantu calon guru dalam memahami dan menerapkan analisis CTT dengan lebih mudah. Adapun fitur utama yang dikembangkan meliputi: (1) antarmuka pengguna (*user interface*) yang ramah dan interaktif, (2) dukungan penggunaan bahasa Indonesia, (3) interpretasi otomatis terhadap hasil analisis, serta (4) pemberian saran atau rekomendasi yang kontekstual pada kondisi-kondisi tertentu dalam proses analisis.

2. Metode Penelitian

Jenis penelitian dalam penelitian ini menggunakan pendekatan *Research and Development* (R&D). Fokus pada penelitian ini adalah pengembangan *software* untuk uji normalitas data. *Software* yang dikembangkan akan menghasilkan output berupa hasil analisis data beserta interpretasinya. Model pengembangan *software* yang digunakan yaitu model SDLC dengan lima aktivitas utama yaitu : (1) perencanaan (planning); (2) Analisis kebutuhan, yang berfokus pada identifikasi kebutuhan pengguna dan spesifikasi sistem; (3) Perancangan (design), meliputi pembuatan struktur antarmuka, rancangan algoritma, serta model data; (4) Implementasi, yaitu tahap penerapan rancangan ke dalam bentuk kode program yang dapat dijalankan; dan (5) system, untuk memastikan perangkat lunak berfungsi sesuai dengan tujuan dan menghasilkan keluaran yang valid (Pressman, 2015; Tilley & Rosenblatt, 2016). Model SDLC sering disebut juga dengan linear sequensial model atau classic life cycle. Pada model ini pelaksanaan dari suatu sistem dilakukan secara berurutan, yaitu jika langkah pertama belum selesai maka langkah kedua tidak dapat dikerjakan, jika langkah kedua belum selesai dikerjakan maka langkah ketiga belum dapat dikerjakan, demikian seterusnya sampai semua tahapan dalam proses pengembangan sistem selesai dikerjakan.

3. Hasil dan Pembahasan

3.1 Hasil

Bagian ini menjelaskan proses pengemasan konten CTT ke dalam perangkat lunak yang dikembangkan berdasarkan model *Software Development Life Cycle* (SDLC). Setiap

tahapan dalam proses pengembangan dijabarkan secara sistematis untuk menggambarkan hasil dan implementasi yang diperoleh pada masing-masing tahap.

3.1.1 Prosedur pengembangan

a. *Planning*,

Tahapan ini menghasilkan sejumlah keputusan penting yang berkaitan dengan arah dan tujuan pengembangan, pemilihan bahasa pemrograman, serta penetapan nama perangkat lunak yang dikembangkan. Tujuan pengembangan beserta fitur-fitur utama dari paket ini telah diuraikan secara rinci pada bagian latar belakang. Perangkat lunak dikembangkan dengan menggunakan bahasa pemrograman R (R Core Team, 2023) yang diimplementasikan melalui lingkungan pengembangan RStudio (Posit team, 2023). Produk yang dihasilkan dalam penelitian ini diberi nama 3S CTT (*Software Statistika Sederhana seri Classical Theory Test*).

b. *Analysist*,

- 1) Berdasarkan hasil analisis kebutuhan, ditetapkan bahwa terdapat tiga model *Classical Test Theory* (CTT) yang dapat dianalisis menggunakan perangkat lunak **3S CTT**, yaitu: (1) model dikotomis, (2) model politomis untuk soal esai, dan (3) model politomis untuk instrumen angket.
- 2) Prosedur analisis CTT dalam perangkat lunak ini disusun berdasarkan beberapa *layer* utama yang merepresentasikan alur kerja pengguna. Empat *layer* antarmuka pengguna (*user interface*) yang digunakan dalam **3S CTT** meliputi: *input layer*, *hasil analisis layer*, *interpretasi per butir layer*, dan *penskoran layer*.
- 3) Rincian hasil analisis kebutuhan yang menjadi dasar dalam pengembangan antarmuka pengguna dan komponen analisis CTT disajikan secara lengkap pada **Tabel 1**.

Tabel 1. Analisa kebutuhan paket R yang digunakan untuk pendukung pengembangan

| Komponen | paket | Fungsi paket |
|-----------------|---------------------------------------|---|
| Interface | shiny(Perrier et al., 2023), | untuk membangun user interface yang terdiri dari bagian ui.R dan server.R |
| | | |
| | shinyWidget(Perrier et al., 2023) | untuk membuat button, radio button, dan radio group button. |
| | shinythemes(Chang, 2021) | Untuk melengkapi package Shiny |
| | Ggplot2(Wickham et al., 2007) | Untuk menampilkan gambar perbandingan daya pembeda |
| | DT(Xie et al., 2023) | untuk menampilkan tabel. |
| Perhitungan CTT | gt(Iannone et al., 2022) | untuk membuat tampilan tabel. |
| | CTT | untuk analisis butir instrumen. |
| | psych(William Revelle, 2023) | untuk menentukan uji reliabilitas |
| Input data | itemanalysis(Zopluoglu, 2018) | untuk analisis butir instrumen |
| | readxl(Wickham & Bryan, 2023) | untuk input data bereskstensi xlsx |
| Laporan | Openxlsx (Schauburger & Walker, 2023) | Untuk membuat File laporan Hasil Analisis CTT |

c. *Design*,

Pada tahap ini dilakukan perancangan *user interface* (antarmuka pengguna) yang berfungsi sebagai media interaksi utama antara pengguna dan sistem. Desain antarmuka disusun berdasarkan jumlah *layer* yang telah ditetapkan pada tahap sebelumnya, sehingga setiap komponen tampilan selaras dengan fungsi analisis yang diinginkan. Gambaran umum mengenai rancangan antarmuka serta elemen visual dari produk dapat dilihat pada bagian penjelasan produk.

d. *Implementation*,

1) pengkodean

Bahasa R digunakan untuk pengkodean. Proses pengkodean dilakukan dengan bantuan software R Studio (Posit team, 2023). Hasil pengkodean terdiri dari satu file utama yaitu app.R yang berisi file front-end memuat kode-kode pengontrol user interface. (2) file back-end memuat kode-kode dari paket *CTT*, *Itemanalysis*, *psych*, *readxl*, *diagram*, dan *stats*.

2) proses uji coba

Uji coba dilakukan dengan membandingkan hasil output *3SCTT* dengan output *itemanalysis* dan *CTT*. Uji coba dilakukan berkali-kali dengan menggunakan data dengan format dikotomus dan politomus. Hasil ujicoba dengan para pakar diperoleh hasil pengujian seperti pada tabel 2.

Table 2. hasil uji coba paket *3SCTT*

| Komponens | Fungsionalitas <i>3SCTT</i> |
|---------------------------|-----------------------------|
| Reliabilitas | ✓ |
| Parameter butir dikotomus | ✓✓ |
| Parameter butir politomus | ✓✓✓ |

Tanda ✓ menunjukkan hasil yang sama dengan hasil analisis paket *psych* secara manual

Tanda ✓✓ menunjukkan hasil yang sama dengan hasil analisis paket *CTT* secara manual

Tanda ✓✓✓ menunjukkan hasil yang sama dengan hasil analisis paket *Itemanalysis* secara manual

e. *System*

Sistem pada pengembangan software dikemas dalam bentuk web app. Penyedia layanan yang digunakan ShinyApp.io. Paket dapat dilihat pada link <https://3sctt.shinyapps.io/3sctt/>

3.1.2 Ilustrasi penggunaan Software

3.1.2.1 Akses Link

Masuk ke link <https://3sctt.shinyapps.io/3sctt/> untuk dapat menggunakan *3SCTT*.

3.1.2.2 Tentang Software

3S-CTT Input Data Hasil Analisis Analisis Tiap Butir Penskoran dan Download Tentang Software

TENTANG 3S-CTT
Software Statistika Sederhana - Seri Classical Theory Test (3S-CTT) merupakan Software yang dapat digunakan untuk 1) Analisis Butir yang didasarkan pada Teori Tes Klasik, 2) Untuk Analisis Butir Tes Pilihan Ganda, Essay dan Non Test. Kelebihan dari software ini yaitu hanya dengan 1 klik tombol ANALISIS, Secara otomatis akan diperoleh (a) Hasil Analisis,(b) Interpretasi Hasil, (c) Download Hasil Analisis CTT.

KONTRIBUTOR

1. Hari Purnomo Susanto (STKIP PGRI Pacitan): Review software, Konsep Teori Tes Klasik, Koding dan Desain software Pengajar Evaluasi dan Analisis BUTir
2. Fathurrahman (Universitas Pendidikan Muhammadiyah Sorong, Papua Barat Daya): Pakar assessment dan Evaluasi Pendidikan Fisika
3. Mega Isvandiana Purnamasari (STKIP PGRI Pacitan): Praktisi Psikologi
4. Khoirul Qudsiyah (STKIP PGRI Pacitan): Pengajar Evaluasi dan Analisis BUTir
5. Johri Sabariyati(Universitas Muhammadiyah Mataram, Mataram NTB): Pakar assessment dan Evaluasi Pendidikan Fisika
6. Jero Budi Darmayasa (universitas Borneo Tarakan): Praktisi pada Analisis Butir Matematika
7. Suciati (universitas Borneo Tarakan): Pakar assessment dan Evaluasi Pendidikan Matematika
8. Eko Wahyunanto Prihono (Institut Agama Islam Negeri Ambon): Pakar Assessment dan Evaluasi Pendidikan
9. Mobinta Kusuma (Universitas Pancasakti Tegal): Assessment dan Evaluasi Pendidikan

Gambar 1. Informasi Tentang Software

Pada layer ini berisi tentang informasi software. Infomasi tersebut berkaitan dengan Fungsionalitas Software, kontributor, dan paket yang digunakan.

3.1.2.3 Persiapan data

Data yang digunakan harus di input pada file exel dengan ekstensi xlsx. Selain ekstensi tersebut tidak dapat digunakan.

1) Data dikotomus

Gambar 2. Seting data dikotomus

Setting data untuk data dikotomus dapat dilakukan seperti pada gambar 2. Kunci dapat dituliskan pada baris pertama atau pada gambar pada baris kedua setelah nama butir. Pada ilustrasi gambar ditampilkan butir yang dianalisis ada 15.

2) Data polytomus

Gambar 3. Setting data polytomus

Setting data polytomus dapat dilakukan seperti pada gambar 3. Pada gambar menunjukkan banyak butir ada 28.

3.1.2.4 Halaman Input Data

Gambar 4. Layer Input data

Gambar 4 menampilkan **layer input data**, yang berfungsi sebagai tahap awal dalam proses analisis menggunakan **3S CTT**. Pada bagian ini tersedia fitur *Upload Data* yang memungkinkan pengguna untuk mengunggah file berekstensi .xlsx berisi data yang akan dianalisis. Selanjutnya, pengguna dapat mengatur **opsi pilihan** atau **penskoran** sesuai dengan karakteristik instrumen yang digunakan. Sebagai contoh, pada instrumen berbentuk pilihan ganda dengan empat alternatif jawaban (A, B, C, dan D), kolom opsi perlu diisi dengan angka “4”. Sementara itu, untuk instrumen politomus, misalnya dengan sistem penskoran 1, 2, 3, 4, dan 5—maka nilai yang diinput adalah “5”.

Hasil uji coba menunjukkan bahwa **3S CTT** mampu menangani data politomus dengan tingkat penskoran hingga sembilan kategori (1–9), sedangkan untuk tingkat di atasnya belum dilakukan pengujian lebih lanjut. Fitur terakhir pada *layer* ini adalah tombol “**Analisis**”, yang berfungsi untuk memulai proses perhitungan. Pada kasus politomus, akan muncul tambahan fitur untuk memilih jenis instrumen, yaitu *tes* atau *skala Likert*. Setelah tombol “Analisis” ditekan, sistem secara otomatis menjalankan prosedur analisis CTT dan menampilkan hasilnya pada *layer-layer* berikutnya secara terintegrasi.

3.1.2.5 Halaman Hasil analisis CTT

Pada bagian ini terdapat empat fitur yaitu fitur Hasil analisis Reliabilitas, Interpretasi Hasil analisis Reliabilitas, Hasil analisis Butir, dan Interpretasi hasil analisi setiap butir.

3S-CTT Input Data Hasil Analisis Analisis Tiap Butir Penskoran dan Download Tentang Software

1. ANALISIS RELIABILITAS

ANALISIS RELIABILITAS

Enam Indikator untuk Mengukur Kualitas Internal Instrumen

| Alpha | Uni | CFI | ECV | EVR | MAP |
|-------|------|------|------|------|------|
| 0.68 | 0.38 | 0.83 | 0.73 | 1.89 | 0.02 |

2. INTERPRETASI HASIL ANALISI RELIABILITAS

Interpretasi Hasil Analisis Reliabilitas tes pada tabel diatas ditampilkan sebagai berikut.

1. Cronbach's Alpha (α): 0.68
 - Kategori: **Cukup** (Dapat diterima untuk eksplorasi awal, tetapi tidak direkomendasikan untuk pengambilan keputusan penting).
2. Unidimensionality (Uni): 0.38
 - Skala **multidimensional** (mengukur lebih dari satu konstruk). Pertimbangkan analisis faktor konfirmatori (CFA).
3. Explained Common Variance (ECV): 0.73
 - Faktor umum dominan (Sebagian besar varian dijelaskan oleh konstruk utama).
4. Eigenvalue Ratio (EVR): 1.89
 - Faktor umum lebih kuat daripada faktor spesifik (mendukung unidimensionalitas).
5. Mean Absolute Partial Correlation (MAP): 0.02
 - Noise rendah (korelasi residual kecil, model faktor cocok).

Gambar 5a. Layer hasil analisis reliabilitas

Pada gambar 5a ditampilkan hasil analisis reliabilitas dan interpretasinya.

3. HASIL ANALISI BUTIR

| Nomor Butir | Analisis Butir Soal dengan Teori Tes Klasik | | | | | | | |
|-------------|---|-----------------|-----------|----------|--------------------------|-------------|--------------------------------------|-------------------------------|
| | Tingkat Kesulitan | | Daya Beda | | Kualitas Butir dan Saran | | Reliabilitas Tes Jika Butir Di Hapus | Distraktor yang Tidak Efektif |
| KoefTK | Kriteria | Keterangan | KoefDB | Kriteria | Keterangan | | | |
| X1 | 0.37 | 0.3<=KoefTK<0.7 | Sedang | 0.288 | 0.2<=KoefDB<0.4 | Baik | Sangat Baik | 0.495 |
| X2 | 0.18 | KoefTK<0.3 | Sulit | 0.437 | KoefDB>0.7 | Sangat Baik | - | 0.469 |
| X3 | 0.45 | 0.3<=KoefTK<0.7 | Sedang | -0.064 | KoefDB<0.2 | Kurang Baik | Direvisi atau Dibuang | 0.581 |
| X4 | 0.28 | KoefTK<0.3 | Sulit | 0.153 | KoefDB<0.2 | Kurang Baik | Dibuang | 0.528 |
| X5 | 0.23 | KoefTK<0.3 | Sulit | 0.330 | 0.2<=KoefDB<0.4 | Baik | Perlu Revisi Isi Soal | 0.488 |
| X6 | 0.36 | 0.3<=KoefTK<0.7 | Sedang | 0.266 | 0.2<=KoefDB<0.4 | Baik | Sangat Baik | 0.501 |
| X7 | 0.11 | KoefTK<0.3 | Sulit | 0.303 | 0.2<=KoefDB<0.4 | Baik | Perlu Revisi Isi Soal | 0.502 |
| X8 | 0.14 | KoefTK<0.3 | Sulit | 0.376 | 0.2<=KoefDB<0.4 | Baik | Perlu Revisi Isi Soal | 0.486 |
| X9 | 0.16 | KoefTK<0.3 | Sulit | 0.323 | 0.2<=KoefDB<0.4 | Baik | Perlu Revisi Isi Soal | 0.494 |
| X10 | 0.21 | KoefTK<0.3 | Sulit | 0.208 | 0.2<=KoefDB<0.4 | Baik | Perlu Revisi Isi Soal | 0.515 |
| X11 | 0.38 | 0.3<=KoefTK<0.7 | Sedang | -0.198 | KoefDB<0.2 | Kurang Baik | Direvisi atau Dibuang | 0.608 |
| X12 | 0.08 | KoefTK<0.3 | Sulit | 0.157 | KoefDB<0.2 | Kurang Baik | Dibuang | 0.526 |
| X13 | 0.14 | KoefTK<0.3 | Sulit | 0.391 | 0.2<=KoefDB<0.4 | Baik | Perlu Revisi Isi Soal | 0.483 |
| X14 | 0.22 | KoefTK<0.3 | Sulit | 0.215 | 0.2<=KoefDB<0.4 | Baik | Perlu Revisi Isi Soal | 0.514 |
| X15 | 0.23 | KoefTK<0.3 | Sulit | 0.020 | KoefDB<0.2 | Kurang Baik | Dibuang | 0.554 |

Gambar 5b. Layer hasil analisis butir

Gambar 5b. berada pada layer kedua yang menunjukkan hasil analisis butir. Pada bagian ini berisi tentang informasi parameter butir, reliabilitas jika butir tertentu dihapus dan informasi kualitas distraktor.

4. INTERPRETASI HASIL ANALISIS BUTIR

Berikut merupakan interpretasi dari hasil analisis butir untuk setiap butir.

Interpretasi

Butir nomor X1 memiliki koefisien tingkat kesulitan (KoefTK) sebesar 0.37, yang berada dalam kategori $0.3 <= \text{KoefTK} < 0.7$. Dengan demikian, butir ini tergolong sebagai soal yang sedang. Sementara itu, daya beda butir ini memiliki koefisien (KoefDB) sebesar 0.288, yang termasuk dalam kategori $0.2 <= \text{KoefDB} < 0.4$, sehingga kemampuan butir dalam membedakan peserta berkemampuan tinggi dan rendah dinilai baik. Secara keseluruhan, kualitas butir ini dinilai Sangat Baik. Apabila butir ini dihapus, maka reliabilitas tes akan menjadi sebesar 0.495. Seluruh distraktor pada butir ini berfungsi dengan baik dan mampu mengalihkan perhatian peserta yang kurang memahami materi.

Butir nomor X2 memiliki koefisien tingkat kesulitan (KoefTK) sebesar 0.18, yang berada dalam kategori $\text{KoefTK} < 0.3$. Dengan demikian, butir ini tergolong sebagai soal yang sulit. Sementara itu, daya beda butir ini memiliki koefisien (KoefDB) sebesar 0.437, yang termasuk dalam kategori $\text{KoefDB} < 0.7$, sehingga kemampuan butir dalam membedakan peserta berkemampuan tinggi dan rendah dinilai sangat baik. Apabila butir ini dihapus, maka reliabilitas tes akan menjadi sebesar 0.469. Namun, terdapat distraktor yang kurang efektif yaitu opsi B,C,D. Distraktor tersebut tidak banyak dipilih oleh peserta yang menjawab salah.

Butir nomor X3 memiliki koefisien tingkat kesulitan (KoefTK) sebesar 0.45, yang berada dalam kategori $0.3 <= \text{KoefTK} < 0.7$. Dengan demikian, butir ini tergolong sebagai soal yang sedang. Sementara itu, daya beda butir ini memiliki koefisien (KoefDB) sebesar -0.064, yang termasuk dalam kategori $\text{KoefDB} < 0.2$, sehingga kemampuan butir dalam membedakan peserta berkemampuan tinggi dan rendah dinilai kurang baik. Secara keseluruhan, kualitas butir ini dinilai Direvisi atau Dibuang. Apabila butir ini dihapus, maka reliabilitas tes akan menjadi sebesar 0.581.

Butir nomor X4 memiliki koefisien tingkat kesulitan (KoefTK) sebesar 0.28, yang berada dalam kategori $\text{KoefTK} < 0.3$. Dengan demikian, butir ini tergolong sebagai soal yang sulit. Sementara itu, daya beda butir ini memiliki koefisien (KoefDB) sebesar 0.153, yang termasuk dalam kategori $0.2 <= \text{KoefDB} < 0.4$, sehingga kemampuan butir dalam membedakan peserta berkemampuan tinggi dan rendah dinilai kurang baik. Secara keseluruhan, kualitas butir ini dinilai Dibuang. Apabila butir ini dihapus, maka reliabilitas tes akan menjadi sebesar 0.528.

Butir nomor X5 memiliki koefisien tingkat kesulitan (KoefTK) sebesar 0.23, yang berada dalam kategori $\text{KoefTK} < 0.3$. Dengan demikian, butir ini tergolong sebagai soal yang sulit. Sementara itu, daya beda butir ini memiliki koefisien (KoefDB) sebesar 0.33, yang termasuk dalam kategori $0.2 <= \text{KoefDB} < 0.4$, sehingga kemampuan butir dalam membedakan peserta berkemampuan tinggi dan rendah dinilai baik. Secara keseluruhan, kualitas butir ini dinilai Sangat Baik. Apabila butir ini dihapus, maka reliabilitas tes akan menjadi sebesar 0.488.

Butir nomor X6 memiliki koefisien tingkat kesulitan (KoefTK) sebesar 0.36, yang berada dalam kategori $0.3 <= \text{KoefTK} < 0.7$. Dengan demikian, butir ini tergolong sebagai soal yang sedang. Sementara itu, daya beda butir ini memiliki koefisien (KoefDB) sebesar 0.266, yang termasuk dalam kategori $0.2 <= \text{KoefDB} < 0.4$, sehingga kemampuan butir dalam membedakan peserta berkemampuan tinggi dan rendah dinilai baik. Secara keseluruhan, kualitas butir ini dinilai Sangat Baik. Apabila butir ini dihapus, maka reliabilitas tes akan menjadi sebesar 0.501.

Butir nomor X7 memiliki koefisien tingkat kesulitan (KoefTK) sebesar 0.11, yang berada dalam kategori $\text{KoefTK} < 0.3$. Dengan demikian, butir ini tergolong sebagai soal yang sulit. Sementara itu, daya beda butir ini memiliki koefisien (KoefDB) sebesar 0.303, yang termasuk dalam kategori $0.2 <= \text{KoefDB} < 0.4$, sehingga kemampuan butir dalam membedakan peserta berkemampuan tinggi dan rendah dinilai baik. Secara keseluruhan, kualitas butir ini dinilai Sangat Baik. Apabila butir ini dihapus, maka reliabilitas tes akan menjadi sebesar 0.502.

Butir nomor X8 memiliki koefisien tingkat kesulitan (KoefTK) sebesar 0.14, yang berada dalam kategori $\text{KoefTK} < 0.3$. Dengan demikian, butir ini tergolong sebagai soal yang sulit. Sementara itu, daya beda butir ini memiliki koefisien (KoefDB) sebesar 0.376, yang termasuk dalam kategori $0.2 <= \text{KoefDB} < 0.4$, sehingga kemampuan butir dalam membedakan peserta berkemampuan tinggi dan rendah dinilai baik. Secara keseluruhan, kualitas butir ini dinilai Sangat Baik. Apabila butir ini dihapus, maka reliabilitas tes akan menjadi sebesar 0.486.

Butir nomor X9 memiliki koefisien tingkat kesulitan (KoefTK) sebesar 0.16, yang berada dalam kategori $\text{KoefTK} < 0.3$. Dengan demikian, butir ini tergolong sebagai soal yang sulit. Sementara itu, daya beda butir ini memiliki koefisien (KoefDB) sebesar 0.323, yang termasuk dalam kategori $0.2 <= \text{KoefDB} < 0.4$, sehingga kemampuan butir dalam membedakan peserta berkemampuan tinggi dan rendah dinilai baik. Secara keseluruhan, kualitas butir ini dinilai Sangat Baik. Apabila butir ini dihapus, maka reliabilitas tes akan menjadi sebesar 0.494.

Butir nomor X10 memiliki koefisien tingkat kesulitan (KoefTK) sebesar 0.21, yang berada dalam kategori $\text{KoefTK} < 0.3$. Dengan demikian, butir ini tergolong sebagai soal yang sulit. Sementara itu, daya beda butir ini memiliki koefisien (KoefDB) sebesar 0.208, yang termasuk dalam kategori $0.2 <= \text{KoefDB} < 0.4$, sehingga kemampuan butir dalam membedakan peserta berkemampuan tinggi dan rendah dinilai baik. Secara keseluruhan, kualitas butir ini dinilai Sangat Baik. Apabila butir ini dihapus, maka reliabilitas tes akan menjadi sebesar 0.515.

Butir nomor X11 memiliki koefisien tingkat kesulitan (KoefTK) sebesar 0.38, yang berada dalam kategori $0.3 <= \text{KoefTK} < 0.7$. Dengan demikian, butir ini tergolong sebagai soal yang sedang. Sementara itu, daya beda butir ini memiliki koefisien (KoefDB) sebesar -0.198, yang termasuk dalam kategori $\text{KoefDB} < 0.2$, sehingga kemampuan butir dalam membedakan peserta berkemampuan tinggi dan rendah dinilai baik. Secara keseluruhan, kualitas butir ini dinilai Sangat Baik. Apabila butir ini dihapus, maka reliabilitas tes akan menjadi sebesar 0.608.

Butir nomor X12 memiliki koefisien tingkat kesulitan (KoefTK) sebesar 0.08, yang berada dalam kategori $\text{KoefTK} < 0.3$. Dengan demikian, butir ini tergolong sebagai soal yang sulit. Sementara itu, daya beda butir ini memiliki koefisien (KoefDB) sebesar 0.157, yang termasuk dalam kategori $0.2 <= \text{KoefDB} < 0.4$, sehingga kemampuan butir dalam membedakan peserta berkemampuan tinggi dan rendah dinilai baik. Secara keseluruhan, kualitas butir ini dinilai Sangat Baik. Apabila butir ini dihapus, maka reliabilitas tes akan menjadi sebesar 0.526.

Butir nomor X13 memiliki koefisien tingkat kesulitan (KoefTK) sebesar 0.14, yang berada dalam kategori $\text{KoefTK} < 0.3$. Dengan demikian, butir ini tergolong sebagai soal yang sulit. Sementara itu, daya beda butir ini memiliki koefisien (KoefDB) sebesar 0.391, yang termasuk dalam kategori $0.2 <= \text{KoefDB} < 0.4$, sehingga kemampuan butir dalam membedakan peserta berkemampuan tinggi dan rendah dinilai baik. Secara keseluruhan, kualitas butir ini dinilai Sangat Baik. Apabila butir ini dihapus, maka reliabilitas tes akan menjadi sebesar 0.483.

Butir nomor X14 memiliki koefisien tingkat kesulitan (KoefTK) sebesar 0.22, yang berada dalam kategori $\text{KoefTK} < 0.3$. Dengan demikian, butir ini tergolong sebagai soal yang sulit. Sementara itu, daya beda butir ini memiliki koefisien (KoefDB) sebesar 0.215, yang termasuk dalam kategori $0.2 <= \text{KoefDB} < 0.4$, sehingga kemampuan butir dalam membedakan peserta berkemampuan tinggi dan rendah dinilai baik. Secara keseluruhan, kualitas butir ini dinilai Sangat Baik. Apabila butir ini dihapus, maka reliabilitas tes akan menjadi sebesar 0.514.

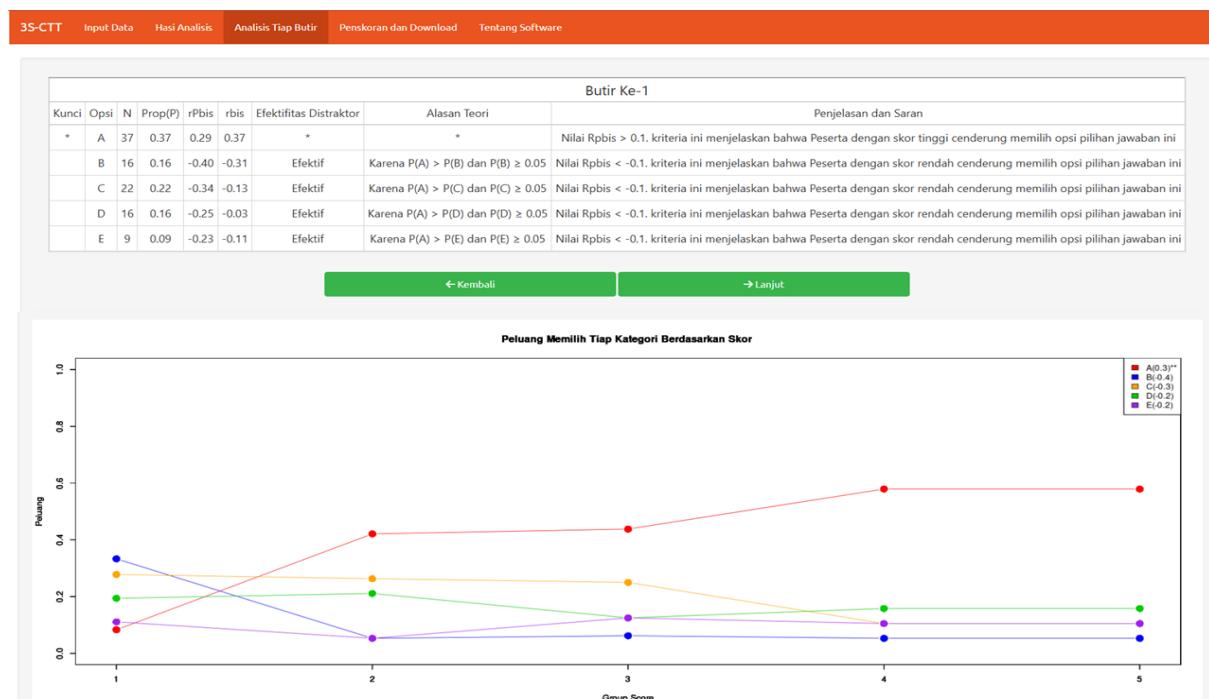
Butir nomor X15 memiliki koefisien tingkat kesulitan (KoefTK) sebesar 0.23, yang berada dalam kategori $\text{KoefTK} < 0.3$. Dengan demikian, butir ini tergolong sebagai soal yang sulit. Sementara itu, daya beda butir ini memiliki koefisien (KoefDB) sebesar 0.020, yang termasuk dalam kategori $\text{KoefDB} < 0.2$, sehingga kemampuan butir dalam membedakan peserta berkemampuan tinggi dan rendah dinilai baik. Secara keseluruhan, kualitas butir ini dinilai Sangat Baik. Apabila butir ini dihapus, maka reliabilitas tes akan menjadi sebesar 0.554.

Gambar 5c. layer hasil analisis interpretasi butir

Bagian ini merupakan interpretasi tentang kualitas secara umum setiap butir. Interpretasi ini didasarkan pada informasi yang ada pada gambar 5b.

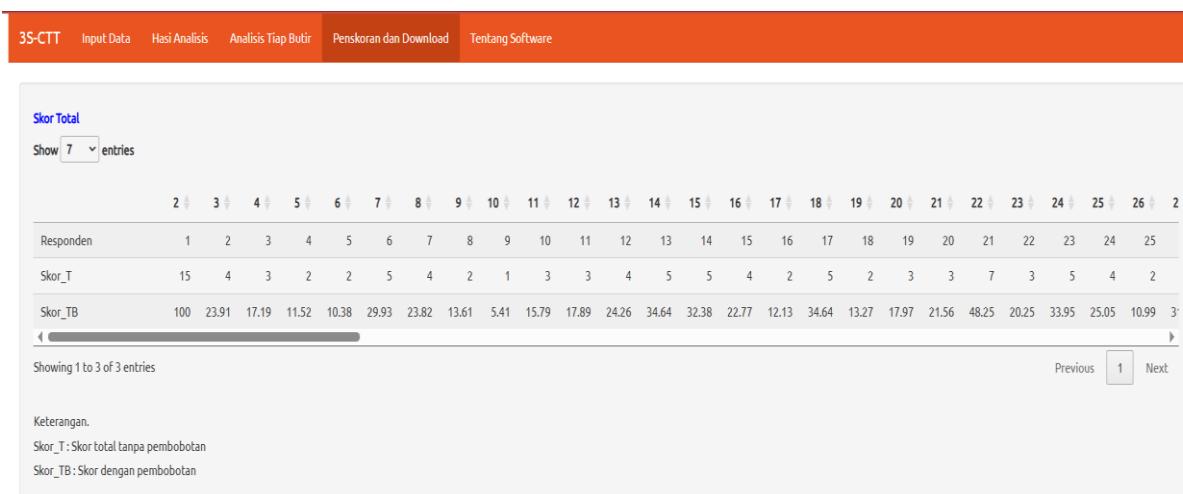
3.1.2.6 Halaman Analisis setiap butir

Pada bagian ini ditampilkan informasi tentang parameter setiap butir beserta distraktornya. Selain hasil, pada bagian ini juga ditampilkan interpretasi kualitas distraktor dan dapat memberikan saran. Bagian ini secara lengkap ditampilkan pada gambar 6. Ilustrasi pada gambar 6 merupakan butir ke-1 dari 15 butir yang digunakan. Untuk melihat informasi butir ke-2 dan sterusnya dapat digunakan tombol lanjut dan Kembali.



Gambar 6. Layer analisis setiap butir

3.1.2.7 Halaman Penskoran dan donwload

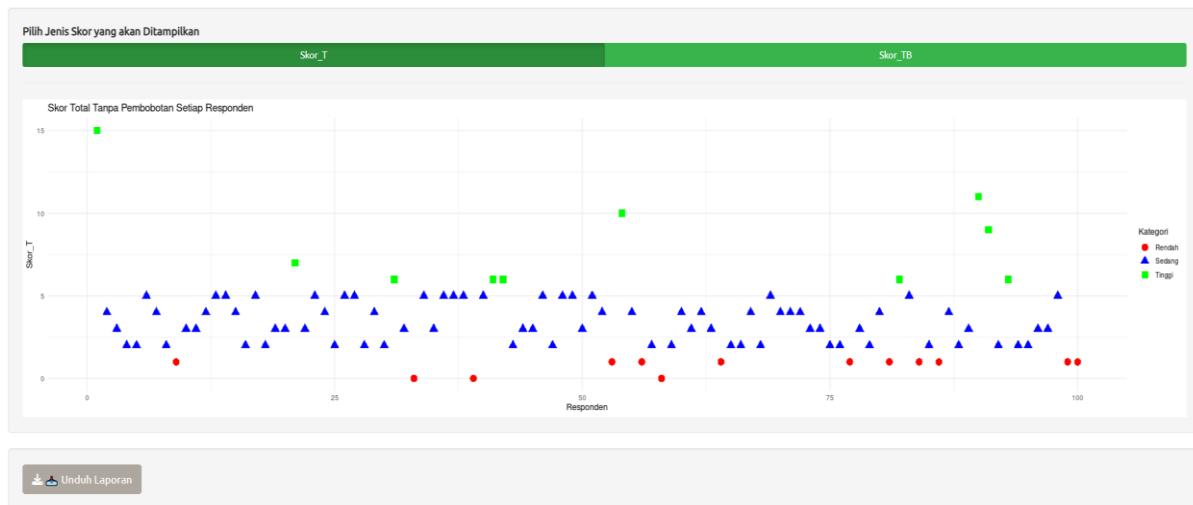


Gambar 7a. Penskoran

Pada bagian ini penskoran data dikotomus dilakukan dengan dua cara yaitu skor total, dan skor bobot. Skor total dilakukan dengan menjumlahkan semua soal benar.

Selanjutnya skor bobot diberikan dengan perhitungan satu dikurangi Tingkat kesulitan. Pembobotan ini dilakukan mengingat Tingkat kesulitan pada konsep CTT ditentukan dengan peluang siswa menjawab benar dibagi skor maksimum. Idealnya butir mudah diberi bobot kecil dan sulit diberi skor besar, maka skor akan berbanding lurus dengan negasi Tingkat kesulitan. Pembobotan ini akan memberikan pembeda nyata yang didasarkan pada Tingkat kesulitan.

Pada bagian ini, jika data yang digunakan politomus, maka akan ditampilkan penskoran berdasarkan total skor dan skala rating.



Gambar 7b. visual hasil skor

Pada bagian ini diberikan bentuk visual berdasarkan kategorisasi empiric. Kategorisasi pada tiga kelompok yaitu siswa dengan skor rendah, sedang dan tinggi.

3.2 Pembahasan

Software 3SCTT dikembangkan sebagai solusi praktis dalam melakukan analisis butir berdasarkan pendekatan Classical Test Theory (CTT). Perangkat lunak ini menawarkan sejumlah keunggulan yang belum tentu tersedia pada aplikasi sejenis, seperti Iteman, SPSS, atau template Excel manual. Salah satu keunggulan utama 3SCTT terletak pada kemudahan proses input data, di mana calon guru tidak perlu melakukan pengaturan teknis yang kompleks, seperti pengkodean variabel atau konversi format file tertentu. Sebagai perbandingan, penggunaan Iteman memerlukan file berformat tetap (.dat) serta konfigurasi awal yang cukup teknis, sementara SPSS menuntut pemahaman statistik yang relatif lebih mendalam. Selain itu, 3SCTT menggunakan antarmuka berbahasa Indonesia, sehingga lebih inklusif bagi pengguna lokal dan dapat mengurangi risiko kesalahan dalam memahami istilah teknis. Keunggulan lainnya adalah sifatnya yang berbasis web, memungkinkan calon guru mengaksesnya tanpa proses instalasi tambahan. Dengan demikian, calon guru dapat menjalankan aplikasi ini secara lintas perangkat dan platform. Seluruh layanan dalam 3SCTT disediakan secara gratis, menjadikannya alat yang sangat terjangkau dan dapat dimanfaatkan tidak hanya oleh calon guru, tetapi juga guru, dosen, maupun peneliti di berbagai jenjang pendidikan.

Dari segi fungsionalitas, 3SCTT dirancang untuk menyajikan hasil analisis butir secara komprehensif, lengkap dengan interpretasi yang mudah dipahami oleh pengguna. Setiap butir dianalisis berdasarkan tingkat kesukaran, daya beda, serta efektivitas distraktor (khusus untuk soal pilihan ganda), dan ditampilkan dengan penjelasan naratif dalam bahasa Indonesia. Selain itu, perangkat lunak ini menyajikan visualisasi sederhana untuk memudahkan identifikasi karakteristik butir, baik yang efektif maupun yang perlu diperbaiki. Proses penskoran dilakukan secara otomatis, baik untuk skor total maupun skor berbobot, yang memungkinkan analisis lebih rinci terhadap performa peserta tes. Tidak hanya terbatas pada soal dikotomus (skor 0–1), 3SCTT juga mendukung analisis terhadap data polytomus, seperti butir esai atau kuesioner berskala, sehingga cocok untuk berbagai bentuk instrumen penilaian.

Salah satu aspek inovatif dari 3SCTT adalah pendekatannya yang humanis dan edukatif. Perangkat lunak ini tidak hanya menyajikan hasil dalam bentuk angka statistik, melainkan juga memberikan penjelasan interpretatif yang selaras dengan teori CTT, sehingga membantu calon guru memahami makna di balik hasil analisis secara konseptual. Setiap hasil analisis, seperti tingkat kesukaran, daya beda, dan efektivitas distraktor, disajikan dengan penjelasan teoretis yang ringkas dan mudah dipahami, sehingga memperkuat pemahaman calon guru terhadap prinsip-prinsip pengukuran dan evaluasi.

Antarmuka yang sederhana dan alur penggunaan yang logis memungkinkan pengguna dari berbagai latar belakang, termasuk yang tidak memiliki keahlian statistik atau psikometri, untuk mengakses dan menggunakan aplikasi ini dengan nyaman. Dengan demikian, 3SCTT berkontribusi dalam meningkatkan literasi asesmen sekaligus membangun pemahaman konseptual tentang kualitas butir secara berkelanjutan.

Bagi calon guru, keberadaan **3SCTT** memiliki implikasi yang sangat penting dalam penguatan kompetensi asesmen dan pengukuran hasil belajar. Selama ini, banyak mahasiswa calon guru mengalami kesulitan dalam memahami dan menerapkan teori-teori pengukuran seperti *Classical Test Theory* (CTT), terutama pada tahap analisis butir soal secara manual. Melalui 3SCTT, proses analisis yang biasanya memerlukan pemahaman statistik mendalam dapat dilakukan secara lebih sederhana, tanpa mengabaikan akurasi hasil. Dengan demikian, calon guru tidak hanya mampu melakukan analisis butir secara praktis, tetapi juga memahami makna teoretis dari setiap hasil analisis yang dihasilkan perangkat lunak tersebut. Hal ini mendukung penguasaan kompetensi pedagogik, khususnya dalam aspek perancangan dan evaluasi pembelajaran berbasis data.

Lebih lanjut, 3SCTT berfungsi sebagai media pembelajaran berbasis teknologi yang dapat menumbuhkan *assessment literacy* pada calon guru. Fitur interpretasi hasil analisis dalam bahasa Indonesia yang mudah dipahami membantu mahasiswa pendidikan untuk mengenali konsep-konsep penting seperti tingkat kesukaran, daya beda, dan efektivitas distraktor secara konseptual. Dengan demikian, aplikasi ini tidak hanya digunakan sebagai alat bantu analisis, tetapi juga sebagai sarana belajar mandiri yang memperkuat keterampilan analitis dan reflektif calon guru dalam mengevaluasi instrumen penilaian.

Selain itu, penggunaan 3SCTT dalam konteks pendidikan calon guru juga mendorong penerapan prinsip pengambilan keputusan berbasis data. Melalui praktik analisis butir yang dilakukan dengan aplikasi ini, calon guru belajar untuk mengandalkan data empirik dalam menilai kualitas instrumen, bukan hanya berdasarkan intuisi atau persepsi subjektif. Pendekatan ini penting untuk membangun profesionalisme guru masa

depan yang memiliki kemampuan evaluatif dan diagnostik yang kuat dalam menilai efektivitas pembelajaran serta capaian peserta didik.

Dari sisi pedagogis, 3SCTT juga dapat dimanfaatkan sebagai alat latihan dalam mata kuliah seperti *Evaluasi Pembelajaran*, *Pengukuran Pendidikan*, atau *Pengembangan Instrumen Penilaian*. Dengan antarmuka yang sederhana dan berbasis web, mahasiswa dapat melakukan analisis secara mandiri di berbagai perangkat tanpa hambatan teknis. Hal ini memungkinkan dosen untuk berfokus pada aspek konseptual dan interpretatif dalam pengajaran teori CTT, sementara aspek perhitungannya difasilitasi oleh perangkat lunak. Dengan demikian, pembelajaran menjadi lebih bermakna dan kontekstual karena mahasiswa tidak hanya memahami teori, tetapi juga mampu mengaplikasikannya dalam praktik nyata.

Secara keseluruhan, pengembangan dan pemanfaatan **3SCTT** memiliki implikasi strategis dalam menyiapkan calon guru yang tidak hanya melek teknologi, tetapi juga memiliki kecakapan analitis dalam melakukan evaluasi pembelajaran. Dengan memadukan teori, praktik, dan teknologi, aplikasi ini berkontribusi pada terbentuknya calon pendidik yang reflektif, berbasis data, serta memiliki pemahaman mendalam terhadap prinsip-prinsip pengukuran pendidikan modern.

4. Kesimpulan dan Saran

3SCTT merupakan perangkat lunak berbasis web yang dikembangkan sebagai alat analisis butir praktis, gratis, dan berbahasa Indonesia berdasarkan pendekatan *Classical Test Theory* (CTT). Aplikasi ini dirancang untuk memudahkan pengguna, khususnya calon guru dan pendidik, dalam melakukan analisis butir tanpa memerlukan pengaturan teknis yang kompleks. Selain menyajikan hasil perhitungan statistik, 3SCTT juga dilengkapi dengan interpretasi hasil dan penjelasan teoretis yang bersifat edukatif, sehingga berfungsi tidak hanya sebagai alat analisis, tetapi juga sebagai media pembelajaran CTT yang interaktif. Dukungan terhadap data dikotomus maupun politomus, serta antarmuka yang sederhana dan informatif, menjadikan 3SCTT relevan digunakan oleh berbagai kalangan di bidang pendidikan dan penelitian. Dengan keunggulan tersebut, 3SCTT berkontribusi nyata dalam meningkatkan mutu asesmen pendidikan serta mendorong terbentuknya budaya evaluasi berbasis data.

Ke depan, 3SCTT memiliki potensi pengembangan yang luas. Integrasi dengan platform pembelajaran digital, seperti *Learning Management System* (LMS), akan memperluas fungsionalitasnya dalam konteks asesmen daring. Selain itu, kemampuan untuk mengekspor hasil analisis ke berbagai format file (misalnya PDF, Word, atau Excel) akan meningkatkan kemudahan dokumentasi dan pelaporan hasil evaluasi. Pengembangan lebih lanjut juga dapat mencakup penambahan fitur-fitur cerdas, seperti pemantauan kualitas butir secara longitudinal (*item tracking*) dan rekomendasi revisi otomatis kecerdasan buatan (AI). Dengan pengembangan berkelanjutan, 3SCTT berpotensi menjadi perangkat analisis dan pembelajaran asesmen yang adaptif, inovatif, dan relevan dengan kebutuhan pendidikan modern.

Daftar Pustaka

- Chang, W. (2021). *shinythemes: Themes for Shiny*. <https://cran.r-project.org/package=shinythemes>
- Hayat, B. (2021). Klasika: Program Analisis Item dan Tes dengan Pendekatan Klasik. *Jurnal Pengukuran Psikologi Dan Pendidikan Indonesia (JP3I)*, 10(1), 1–11. <https://doi.org/10.15408/jp3i.v10i1.20551>

- Hu, Z., Lin, L., Wang, Y., & Li, J. (2021). The Integration of Classical Testing Theory and Item Response Theory. *Psychology, 12*.
- Iannone, R., Cheng, J., Schloerke, B., Hughes, E., & Seo, J. (2022). *Package gt* (p. 306).
- Meyer, J. P. (2014). *Applied Measurement with jMetrik*. Routledge. <https://doi.org/10.4324/9780203115190>
- Perrier, V., Meyer, F., & Granjon, D. (2023). *shinyWidgets: Custom Inputs Widgets for Shiny* (R package version 0.7.6). <https://cran.r-project.org/package=shinyWidgets>
- Posit team. (2023). *RStudio: Integrated Development Environment for R*. <http://www.posit.co/>
- Pressman, R. (2015). Rekayasa Perangkat Lunak: Pendekatan Praktisi Buku I. *Jurnal Informa Politeknik Indonusa Surakarta ISSN*, 4(1).
- R Core Team. (2023). *R: A Language and Environment for Statistical Computing*. <https://www.r-project.org/>
- Schauberger, P., & Walker, A. (2023). *openxlsx: Read, Write and Edit xlsx Files*. <https://cran.r-project.org/package=openxlsx>
- Sumaryanta. (2021). Teori Tes Klasik dan Teori Respon Butir: Konsep dan Contoh Penerapannya. In *Cetakan Pertama* (Vol. 15, Issue 2).
- Susanto, H. P., Meifiani, N. I., Purnamasari, M. I., Kusuma, M., Noviantari, I., Sumin, & Prasetyo, T. D. (2023). Introducing Software Statistika Sederhana Seri Analisis Regresi (3SAR). *International Conference on Learning and Education (ICLE)*, 285–292. <https://prosiding.stkippacitan.ac.id/index.php/icle/article/download/36/67>
- Susanto, H. P., Retnawati, H., Abadi, A. M., Haryanto, H., & Djidu, H. (2023). irtawsi: Items Response Theory Analysis with Steps and Interpretation. In *CRAN: Contributed Packages*. <https://doi.org/10.32614/CRAN.package.irtawsi>
- Tilley, S., & Rosenblatt, H. (2016). Systems Analysis and Design, Eleventh Edition. In *A Guide to Medical Computing*.
- Wickham, H., & Bryan, J. (2023). *readxl: Read Excel Files* (R package version 1.4.2). <https://cran.r-project.org/package=readxl>
- Wickham, H., Chang, W., Henry, L., Pedersen, T. L., Takahashi, K., Wilke, C., Woo, K., Yutani, H., Dunnington, D., & van den Brand, T. (2007). *ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics*. In *CRAN: Contributed Packages*. <https://doi.org/10.32614/CRAN.package.ggplot2>
- William Revelle. (2023). *psych: Procedures for Psychological, Psychometric, and Personality Research* (R package version 2.3.3). <https://cran.r-project.org/package=psych>
- Willse, J. T. (2018). *CTT: Classical Test Theory Functions*. In *CRAN: Contributed Packages*. <https://doi.org/10.32614/CRAN.package.CTT>
- Xie, Y., Cheng, J., & Tan, X. (2023). *DT: A Wrapper of the JavaScript Library “DataTables”* (R package version 0.27). <https://cran.r-project.org/package=DT>
- Zopluguoglu, C. (2018). *itemanalysis: Classical Test Theory Item Analysis*. In *CRAN: Contributed Packages*. <https://doi.org/10.32614/CRAN.package.itemanalysis>