

**ANALISIS AKURASI PENGISIAN LEMBAR JAWAB
UJIAN BERBASIS KERTAS MENGGUNAKAN
ALGORITMA *DECISION TREE* C4.5**

**Tesis
untuk memenuhi sebagian persyaratan
mencapai derajat Sarjana S-2 Program Studi
Magister Sistem Informasi**



**Edy Suharto
30000416410002**

**SEKOLAH PASCASARJANA
UNIVERSITAS DIPONEGORO
SEMARANG
2018**

HALAMAN PENGESAHAN

TESIS

Analisis Akurasi Pengisian Lembar Jawab Ujian Berbasis Kertas Menggunakan Algoritma *Decision Tree* C4.5

Oleh:
Edy Suharto
30000416410002

Telah diujikan dan dinyatakan lulus ujian tesis pada tanggal 7 Agustus 2018 oleh tim penguji Program Studi Magister Sistem Informasi Sekolah Pascasarjana Universitas Diponegoro.

Semarang, 7 Agustus 2018

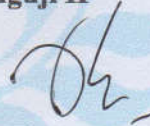
Mengetahui,

Penguji I



Dr. Budi Warsito, S.Si., M.Si.
NIP. 197508241999031003

Penguji II



Djatkiko Endro Suseno, M.Si., Ph.D.
NIP. 197211211998021001

Pembimbing I



Dr. Aris Puji Widodo, S.Si., M.T.
NIP. 197404011999031002

Pembimbing II



Dr. Suryono, S.Si., M.Si.
NIP. 197306301998021001

Mengetahui :

**Dekan Sekolah Pascasarjana
Universitas Diponegoro**



Prof. Dr. Ir. Purwanto, DEA
NIP. 196112281986031004

**Ketua Program Studi
Magister Sistem Informasi**



Dr. Suryono, S.Si., M.Si.
NIP. 197306301998021001

PERNYATAAN

Dengan ini saya menyatakan bahwa dalam tesis ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar akademik di suatu perguruan tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Semarang, 7 Agustus 2018



**PERNYATAAN PERSETUJUAN
PUBLIKASI TESIS UNTUK KEPENTINGAN AKADEMIS**

Sebagai civitas akademik Universitas Diponegoro, saya yang bertanda tangan di bawah ini :

Nama : Edy Suharto
NIM : 30000416410002
Program Studi : Magister Sistem Informasi
Program : Pascasarjana
Jenis Karya : Tesis

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Diponegoro Hak Bebas Royalti Noneksklusif atas karya ilmiah saya yang berjudul :

**Analisis Akurasi Pengisian Lembar Jawab Ujian Berbasis Kertas
Menggunakan Algoritma *Decision Tree* C4.5**

beserta perangkat yang ada. Dengan Hak Bebas Royalti Noneksklusif ini Magister Sistem Informasi Pascasarjana Universitas Diponegoro berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (*database*) merawat, dan mempublikasikan tesis saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Dibuat di : Semarang
Pada tanggal : 7 Agustus 2018

Yang menyatakan



NIM. 30000416410002

KATA PENGANTAR

Puji syukur ke hadirat Allah S.W.T. penulis sampaikan karena atas rahmat-Nya sehingga penulis dapat menyelesaikan tesis dengan judul “Analisis Akurasi Pengisian Lembar Jawab Ujian Berbasis Kertas Menggunakan Algoritma *Decision Tree C4.5*”. Tesis ini disusun untuk memenuhi salah satu persyaratan dalam memperoleh gelar sarjana S-2 pada Program Pascasarjana Magister Sistem Informasi Universitas Diponegoro.

Dalam penyusunan tesis ini, penulis telah mendapat bimbingan dan saran dari berbagai pihak. Untuk itu penulis mengucapkan terima kasih yang sebesar-besarnya kepada :

1. Bapak Prof. Dr. Ir. Purwanto, DEA, selaku Direktur Sekolah Pascasarjana Universitas Diponegoro.
2. Bapak Dr. Suryono, S.Si, M.Si., selaku Ketua Program Studi Magister Sistem Informasi Universitas Diponegoro dan selaku Pembimbing Kedua yang telah mengarahkan penulis dalam menyelesaikan tesis ini.
3. Bapak Dr. Aris Puji Widodo, S.Si., M.T., selaku Pembimbing Pertama yang telah mengarahkan penulis dalam menyelesaikan tesis ini.
4. Semua pihak yang telah memberikan dukungan dalam penelitian ini.

Penulis telah berusaha maksimal dalam penulisan tesis ini, namun kritik, saran dan sumbangan ide dari semua pihak masih sangat diharapkan untuk kesempurnaan. Akhir kata penulis berharap semoga tesis ini dapat bermanfaat dan memberikan wawasan tambahan khususnya bagi civitas akademika Program Pascasarjana Magister Sistem Informasi dan bagi masyarakat pada umumnya.

Semarang, 7 Agustus 2018

Edy Suharto

DAFTAR ISI

	Halaman
Halaman Judul.....	i
Halaman Pengesahan	ii
Pernyataan	iii
Pernyataan Publikasi	iv
Kata Pengantar	v
Daftar Isi.....	vi
Daftar Gambar.....	viii
Daftar Tabel	ix
Daftar Lampiran	x
Daftar Arti Lambang dan Singkatan	xi
Abstrak.....	xii
<i>Abstract</i>	xiii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Tujuan Penelitian	3
1.3 Manfaat Penelitian	4
BAB II TINJAUAN PUSTAKA DAN DASAR TEORI	5
2.1 Tinjauan Pustaka.....	5
2.2 Dasar Teori.....	6
2.2.1 Teori Pengembangan Sistem Informasi	6
2.2.2 Teori <i>Data Mining</i>	7
2.2.3 Algoritma <i>Decision Tree C4.5</i>	9
2.2.4 Teori Pengujian Data	11
BAB III METODE PENELITIAN	13
3.1 Bahan dan Alat Penelitian.....	13
3.2 Prosedur Penelitian	13
3.3 Kerangka Sistem Informasi	15
BAB IV HASIL PENELITIAN DAN PEMBAHASAN	18
4.1 Hasil Penelitian	18
4.1.1 Sistem <i>Preprocessing</i>	18
4.1.2 Klasifikasi	29
4.2 Pembahasan	32
4.2.1 Analisis dan Validasi	32
4.2.2 Evaluasi	38
BAB V KESIMPULAN DAN SARAN.....	40
5.1 Kesimpulan	40
5.2 Saran	40

DAFTAR PUSTAKA	42
LAMPIRAN.....	44

DAFTAR GAMBAR

	Halaman
Gambar 2.1 Siklus <i>Data Mining</i>	8
Gambar 3.1 Prosedur Penelitian	14
Gambar 3.2 Kerangka Sistem Informasi	16
Gambar 4.1 ERD Sistem <i>Preprocessing</i>	19
Gambar 4.2 DCD Sistem <i>Preprocessing</i>	20
Gambar 4.3 DFD Level 1 Sistem <i>Preprocessing</i>	21
Gambar 4.4 DFD Level 2 Pembacaan Masukan	21
Gambar 4.5 DFD Level 2 Konsolidasi	21
Gambar 4.6 Desain Antarmuka Sistem <i>Preprocessing</i>	26
Gambar 4.7 Tampilan Program Web Sistem <i>Preprocessing</i>	27
Gambar 4.8 File CSV Keluaran Program	28
Gambar 4.9 Tampilan Aplikasi WEKA Memuat Data	30
Gambar 4.10 Proses Klasifikasi dengan Uji <i>Cross Validation</i>	31
Gambar 4.11 Proses Klasifikasi dengan Uji <i>Percentage Split</i>	31
Gambar 4.12 Pohon Keputusan Hasil Uji <i>Cross Validation</i>	32
Gambar 4.13 Pohon Keputusan Hasil Uji <i>Percentage Split</i>	33
Gambar 4.14 Instans dengan Status Nomor dan Kode Soal Bernilai Salah	34

DAFTAR TABEL

	Halaman
Tabel 4.1 Spesifikasi Kebutuhan Perangkat Lunak	18
Tabel 4.2 Daftar Perancangan Struktur Data	23
Tabel 4.3 Hasil Pengujian Program	28
Tabel 4.4 Komposisi Nilai Atribut.....	29
Tabel 4.5 Perhitungan Entropi Total.....	36
Tabel 4.6 Perhitungan <i>Gain Ratio</i> Total	37
Tabel 4.7 Perhitungan Entropi Status_ nomor Benar	37
Tabel 4.8 Perhitungan <i>Gain Ratio</i> Status_ nomor Benar	38
Tabel 4.9 <i>Confussion Matrix</i> Hasil <i>Cross Validation</i>	39
Tabel 4.10 <i>Confussion Matrix</i> Hasil <i>Percentage Split</i>	39

DAFTAR LAMPIRAN

	Halaman
Data Referensi Peserta	44
Data Kode Soal	54
Data Lembar Jawab	54
Data Hasil Pelabelan	63

DAFTAR ARTI LAMBANG DAN SINGKATAN

Daftar Arti Lambang

Lambang	Arti Lambang
A	Atribut
\in	Keanggotaan
S	Koleksi sampel
S_v	Koleksi sampel dengan nilai v
p_i	Proporsi sampel ke- i

Daftar Singkatan

Singkatan	Kepanjangan Singkatan
CSV	<i>Comma-separated Values</i>
DCD	<i>Data Context Diagram</i>
DFD	<i>Data Flow Diagram</i>
EDM	<i>Educational Data Mining</i>
ERD	<i>Entity Relationship Diagram</i>
GB	<i>Giga Byte</i>
GHz	<i>Giga Hertz</i>
HTML	<i>Hypertext Markup Language</i>
PHP	<i>PHP Hypertext Preprocessing</i>
SDLC	<i>System Development Life Cycle</i>
SRS	<i>Software Requirements Specification</i>
WEKA	<i>Waikato Environment for Knowledge Analysis</i>

Analisis Akurasi Pengisian Lembar Jawab Ujian Berbasis Kertas Menggunakan Algoritma *Decision Tree* C4.5

ABSTRAK

Dalam penelitian ini telah dilakukan analisis terhadap data hasil pemrosesan lembar jawab ujian berbasis kertas menggunakan teknik *data mining*. Teknik tersebut dapat digunakan untuk melakukan klasifikasi data dan untuk menemukan komponen data penting dalam lembar jawab. Penelitian ini didasarkan pada adanya persoalan kesalahan pengisian data dalam model ujian berbasis kertas yang tidak muncul pada ujian berbasis komputer, sementara akurasi pengisian data ujian penting untuk penjaminan mutu pendidikan. Pada penelitian ini diusulkan metode untuk menganalisis akurasi pengisian data ujian di lembar jawab, yaitu menggunakan Algoritma *Decision Tree* C4.5. Hasil penelitian ini adalah sebuah program berbasis web untuk sistem pra-pemrosesan data dan sebuah model pohon yang mencerminkan kondisi data. Analisis terhadap 374 instans menghasilkan akurasi pengisian lembar jawab sebesar 95,19% dan akurasi klasifikasi sebesar 100%. Hasil penelitian diharapkan dapat memotivasi penyelenggara ujian untuk perbaikan sistem ujian.

Kata kunci : *data mining*, klasifikasi, algoritma C4.5, dan ujian berbasis kertas

Analysis on Accuracy of Filling Out the Answer Sheet in Paper-based Test Using C4.5 Decision Tree Algorithm

ABSTRACT

In this study was conducted analysis on the result of a paper-based test using data mining technique. This technique is applicable to classify data and to learn significant components of answer sheet data. This study was motivated by the problem of incorrect data content in paper-based test nonexistent in computer-based test. Meanwhile, accuracy in filling out the test data is important for quality assurance in education. In this study was proposed a method to analyze the accuracy of answer sheet filling out. To do so, C4.5 Decision Tree Algorithm was used here. The results produced in this study were a web-based program for data preprocessing and a decision tree model representing data behaviour. Based on analysis of 374 instances, there were found that the accuracy of answer sheet filling out reached 95.19% and the accuracy of data classification reached 100%. The results was meant to motivate the test administrator for test improvement.

Keywords : data mining, classification, C4.5 algorithm, and paper-based test