

**KLASIFIKASI KELULUSAN PADA SELEKSI PELAMAR GURU DAN
KARYAWAN MENGGUNAKAN ALGORITMA C4.5**

SKRIPSI

**Diajukan Untuk Memenuhi Salah Satu Syarat Kelulusan
Memperoleh gelar Sarjana Komputer**

Oleh

ARIE KUNCAHYO

2014.69.04.0094



**JURUSAN STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS YUDHARTA PASURUAN**

2018

PERNYATAAN PENULIS

JUDUL : KLASIFIKASI KELULUSAN PADA SELEKSI PELAMAR
GURU DAN KARYAWAN MENGGUNAKAN ALGORITMA
C4.5

NAMA : ARIE KUNCAHYO

NIM : 2014.69.04.0094

“Saya menyatakan dan bertanggung jawab dengan sebenar-benarnya, bahwa skripsi ini hasil karya saya sendiri kecuali cuplikan dan ringkasan yang masing-masing telah saya jelaskan sumbernya. Jika pada waktu selanjutnya ada pihak lain yang meng-klaim bahwa skripsi ini sebagai karyanya, yang disertai dengan bukti-bukti yang cukup, maka saya bersedia untuk dibatalkan gelar Sarjana Komputer saya beserta segala hak dan kewajiban yang melekat pada gelar tersebut “.

Pasuruan, 30 Juli 2018



Arie Kuncahyo
Penulis

PERSETUJUAN SKRIPSI

JUDUL : KLASIFIKASI KELULUSAN PADA SELEKSI PELAMAR
GURU DAN KARYAWAN MENGGUNAKAN ALGORITMA
C4.5

NAMA : ARIE KUNCAHYO

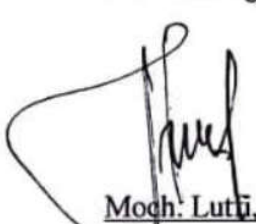
NIM : 2014.69.04.0094

Skripsi ini telah diperiksa dan disetujui

Pasuruan, 30 Juli 2018

Kaprodi,


Lukman Hakim, M.Kom
NIK. Y. 0691101110

Pembimbing,

Moch. Lutfi, M.Kom
NIP. Y. 0691603004

PENGESAHAN SKRIPSI

JUDUL : KLASIFIKASI KELULUSAN PADA SELEKSI PELAMAR GURU DAN KARYAWAN MENGGUNAKAN ALGORITMA C4.5

NAMA : ARIE KUNCAHYO

NIM : 2014.69.04.0094

Skripsi ini telah diujikan dan dipertahankan di depan Dewan Penguji pada Sidang Skripsi tanggal 19 Agustus 2018. Menurut pandangan kami, Skripsi ini memadai dari segi kualitas untuk tujuan penganugerahan gelar Sarjana Komputer (S.Kom)

Pasuruan, 19 Agustus 2018

Ketua Penguji,



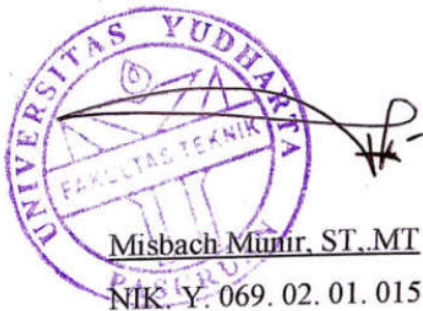
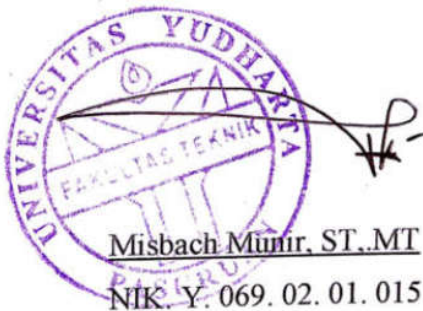
Rahmad Zainul Abidin, M.Kom
NIK. Y. 069. 15. 07. 141

Anggota,



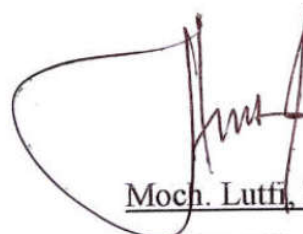
Lukman Hakim, M. Kom
NIK. Y. 069. 11. 01. 110

Dekan Fakultas Teknik,



Misbach Munnir, ST., MT
NIK. Y. 069. 02. 01. 015

Pembimbing,



Moch. Lutfi, M.Kom
NIK. Y. 069. 15. 07. 141

PERSEMBAHAN

Teriring do'a dan rasa syukur kehadiran Allah SWT, penulis persembahkan skripsi ini sebagai ungkapan cinta dan terima kasih kepada:

1. Kedua orang tua tercinta, Ayahanda Sulasmono Arif (Alm) dan Ibunda Endang Priyanti atas ketulusannya dalam mendidik, membesarkan dan membimbing penulis dengan penuh kasih sayang serta keikhlasan didalam iringan do'anya hingga menghantarkan penulis menyelesaikan pendidikan di Universitas Yudharta Pasuruan.
2. Istri tercinta Jurotin Nisa' yang telah memberikan cinta, kasih sayang dan semangat serta dukungan dalam menyelesaikan pendidikan ini.
3. Adikku tersayang Ayu Retno Puji Lestari terima kasih atas doa dan bantuan baik moral maupun material dalam menyelesaikan studi di perguruan tinggi.

ABSTRACT

For increasing the quality of the education is not detached of human resources which own the quality of teachers and employees. The addition of the teachers and employees by releasing the job recruitment provides an opportunity to the applicant to fill in the composition of the work. Data of applicants which exist have a variety of classification. The ways in order to get the applicants that fit the needs of the institutions have been done and those take a right decision so that applicants who passed or not passed can be appropriate. The existing data has the same attributes, therefore it takes a good degree of accuracy and as a result the decision making does not get an error.

This research proposes a method of Algorithm C. 45. This research was conducted in Muhammadiyah senior high school of Sidoarjo. Where the data test is taken from the data of applicants for one year ahead. Accordingly, the level of accuracy c. 45. with the number of folds 5 is 87.89%. While the level of accuracy of the number of folds with a value 20 is 78.33%.

Key : *Algorithm C 4.5, Weka 3.8.1, Classification, Decision Tree*

ABSTRAK

Untuk meningkatkan kualitas pendidikan tidak terlepas dari sumber daya manusia yang dimiliki yaitu kualitas tenaga guru dan karyawan. Penambahan guru dan karyawan dengan membuka lowongan pekerjaan memberikan peluang kepada para pelamar untuk mengisi komposisi pekerjaan. Data pelamar yang ada saat ini memiliki klasifikasi berbagai macam, sudah dilakukan berbagai cara agar bisa mendapatkan pelamar yang sesuai dengan kebutuhan lembaga dan butuh sebuah keputusan yang tepat agar pelamar yang lulus maupun tidak lulus sesuai. Data yang ada memiliki nilai atribut yang sama sehingga dibutuhkan tingkat akurasi yang baik supaya pengambilan keputusan tidak terjadi kesalahan.

Penelitian ini mengusulkan metode Algoritma C.45. Pelaksanaan penelitian ini dilakukan SMA Muhammadiyah Sidoarjo. Dimana data uji adalah data pelamar selama 1 tahun kedepan. Tingkat akurasi C4.5 dengan number of folds 5 memiliki akurasi 87,89 % . sedangkan tingkat akurasi dengan number of folds dengan nilai 20 memiliki nilai akurasi 78,33% .

Kata Kunci : *Algoritma C 4.5, Weka 3.8.1, Klasifikasi, Decision Tree*

KATA PENGANTAR

Puji syukur penulis panjatkan kehadirat Allah SWT yang telah memberikan ilmu pengetahuan, kekuatan dan petunjuk-Nya, sehingga Penulis dapat menyelesaikan skripsi ini yang berjudul: “Klasifikasi Kelulusan Pada Seleksi Pelamar Guru Dan Karyawan Menggunakan Algoritma C4.5”. Sholawat serta salam semoga Allah selalu memberikan Rahmat-Nya kepada Nabi Muhammad SAW, keluarga, para sahabat dan pengikut beliau yang setia.

Penulis menyusun skripsi ini, sebagai bagian dari persyaratan untuk menyelesaikan Pendidikan pada Program Strata Satu (S1) Program Studi Teknik Informatika Universitas Yudharta Pasuruan dan telah penulis selesaikan sesuai dengan rencana. Dalam penulisan skripsi ini, penulis mendapat banyak bantuan dari berbagai pihak khususnya dari dosen pembimbing skripsi, sehingga kesulitan yang dihadapi dapat diselesaikan sesuai dengan harapan. Oleh sebab itu, melalui skripsi ini penulis menyampaikan ucapan terima kasih kepada yang terhormat:

1. KH. Sholeh Baharuddin, selaku pengayom kami selama menempuh pendidikan di Universitas Yudharta Pasuruan.
2. Bapak Dr. M. Saifullah, selaku Rektor Universitas Yudharta Pasuruan
3. Bapak Misbach Munir, ST., MT., selaku Dekan Fakultas Teknik Universitas Yudharta Pasuruan.
4. Bapak Lukman Hakim, M.Kom., selaku Ketua Program Studi Teknik Informatika.
5. Moch. Luthfi, M. Kom selaku Pembimbing yang telah rela meluangkan waktunya untuk mendukung dan membimbing penulis dalam menyelesaikan skripsi ini.
6. Bapak Penguji yang memberikan kritik dan saran pada skripsi ini
7. Dosen-Dosen Program Studi Teknik Informatika yang telah mengajarkan kepada penulis berbagai macam ilmu yang dapat penulis terapkan dalam penulisan skripsi ini
8. Keluarga besar SDN Blarang II yang telah memberikan semangat sehingga memperlancar proses penyusunan skripsi.
9. Kedua Orang Tua yang telah memberikan dukungan moril, semangat dan materiil sehingga memperlancar proses penyusunan skripsi ini.
10. Istriku tercinta yang telah memberikan semangat serta dukungannya hingga terselesainya skripsi ini.
11. Teman-teman seperjuangan Teknik Informatika Angkatan 2014 yang selalu menjadi teman mengejar impian dan mengukir sejarah dalam hidupku, serta menjadi keluarga terbaik selama ini.

12. Semua pihak yang tidak dapat disebutkan satu persatu yang telah berjasa membantu penyelesaian penulisan skripsi ini

Penulis menyadari masih terdapat banyak kekurangan dalam penelitian ini, baik penulisan maupun aplikasinya sendiri. Oleh karena itu penulis mengharapkan saran dan kritik yang dapat membangun skripsi ini lebih baik lagi.

Pasuruan, 19 Juli 2018
Penulis,

Arie Kuncahyo

DAFTAR ISI

PERNYATAAN PENULIS	i
PERSETUJUAN SKRIPSI	ii
PENGESAHAN SKRIPSI	iii
PERSEMBAHAN	iv
ABSTRACT	v
ABSTRAK	vi
KATA PENGANTAR	vii
DAFTAR ISI	ix
DAFTAR GAMBAR	xi
DAFTAR TABEL	xii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan Penelitian.....	3
1.4 Manfaat Penelitian.....	3
1.5 Sistematika Penulisan	3
BAB II TINJAUAN PUSTAKA	5
2.1 Penelitian Terkait	5
2.1.1 Model Penelitian Rusda Wajhillah	5
2.1.2 Model Penelitian Qinbao Song, Martin Shepperd, Xiangru Chena dan Jun Liu	5
2.1.3 Penelitian oleh Siti Mutro	6
2.1.4 Penelitian yang dilakukan oleh Yoso Putra Raharja	6
2.2 Landasan Teori	8
2.1.1 Algoritma C 4.5	8
2.2.2 Pohon Keputusan (Decision Tree)	8
2.2.3 Kelebihan Pohon Keputusan	9
2.2.4 Kekurangan Pohon Keputusan	10

	2.2.5	Arsitektur Pohon Keputusan	10
	2.2.6	Tahapan Algoritma C 4.5	12
	2.2.7	Langkah-Langkah Konstruksi Pohon Keputusan dengan Algoritma C 4.5	13
	2.2.8	Entropy	14
	2.2.9	Gain Ratio	15
	2.2.10	Confusion Matrik	16
	2.3	Kerangka Berfikir	19
BAB III		HASIL DAN PEMBAHASAN	20
	3.1	Analisa Kebutuhan	20
	3.2	Metode Penelitian	20
	3.2.1	Pengumpulan Data.....	23
	3.2.2	Preprocessing	26
	3.3	Metode Yang Diusulkan	26
	3.4	Validasi Dengan Cross Validation	27
	3.5	Confusion Matriks	28
BAB IV		HASIL PENELITIAN DAN PEMBAHASAN	29
	4.1	Hasil Penelitian	29
	4.1.1	Pengumpulan Data.....	29
	4.1.2	Preprocessing dengan Konversi Data	31
	4.1.3	Klasifikasi dengan Algoritma C4.5	32
	4.2	Hasil akurasi data pelamar menggunakan algoritma C4.5	48
	4.3	Hasil klasifikasi dengan aplikasi Weka 3.8.1	48
BAB V		PENUTUP	49
	5.1	Hasil Penelitian	29
	5.2	Saran	48

DAFTAR GAMBAR

Gambar 2.1	Arsitektur Pohon Keputusan.....	11
Gambar 2.2	Kerangka Pemikiran	18
Gambar 3.1	Metode Penelitian	21
Gambar 3.2	Model Yang Diusulkan	27
Gambar 4.1	Node Cabang Pada Atribut IPK	35
Gambar 4.2	Pohon Keputusan Atribut Umur	38
Gambar 4.3	Pohon Keputusan Atribut Pendidikan	41
Gambar 4.4	Pohon Keputusan Atribut Berkas	44
Gambar 4.5	Pohon Keputusan Atribut Wawancara	47
Gambar 4.6	Hasil Klasifikasi dengan Weka 3.8.1	48

DAFTAR TABEL

Tabel 2.1	Tinjauan Pustaka	11
Tabel 3.1	Dataset	25
Tabel 3.2	Value yang Hilang	26
Tabel 4.1	Contoh Dataset	30
Tabel 4.2	Contoh hasil konversi nilai atribut 30 dataset	31
Tabel 4.3	Nilai Entropy Pada Atribut	33
Tabel 4.4	Jumlah Gain pada Atribut	34
Tabel 4.5	Nilai Entropy Selain Atribut IPK	37
Tabel 4.6	Nilai gain masing-masing atribut selain IPK	38
Tabel 4.7	Nilai Entropy Pada Atribut	39
Tabel 4.8	Nilai Gain Pada Atribut	41
Tabel 4.9	Nilai Entropy	42
Tabel 4.10	Perhitungan Gain pada Atribut	44
Tabel 4.11	Nilai entropy pada masing masing atribut	45
Tabel 4.12	Nilai Gain Pada Atribut	46
Tabel 4.13	Perhitungan Akurasi C 4.5	47
Tabel 4.14	Akurasi Menggunakan C 4.5	48