

Sistem Prediksi Pengunduran Diri Calon Mahasiswa Baru Menggunakan Algoritma C45

Supria¹, Lidyawati², Sri Mawarni³

Politeknik Negeri Bengkalis^{1,2,3}

phiya@polbeng.ac.id¹, lidyawati@polbeng.ac.id², srimawarni@polbeng.ac.id³

Abstrak

Polytechnic state of bengkalis is a vocational college located on the island of Bengkalis. The campus has 7 departments, namely shipping, machinery, civil engineering, electrical engineering, informatics engineering, business administration, and maritime. Every year this campus accepts new students. Each department has determined the student quota to be accepted. There are 3 stages that new students must do to become students, namely the registration, selection and re-registration. Prospective students who pass the selection are adjusted to the quota to be accepted. However, there are several prospective students who passed the selection not to re-register, so the predetermined quota was not met. In this research proposed Decision Making System for Resignation of Prospective New Students Using Rule Based C45. This system can be used to predict whether new prospective students will re-register or not. The trial was conducted using prospective students data in 2016 and 2017. From the results of the trials that have been conducted show that the classification of 2016 data has an accuracy rate of 66.53%, 2017 data has an accuracy rate of 71.46% and a combination of 2016 data with 2017 has an accuracy rate of 68.93%.

Keywords : Times New Roman 10 (max 5 words)

1. PENDAHULUAN

Dunia pendidikan terbagi tiga, yaitu pendidikan formal, informal dan nonformal. Dunia pendidikan formal memiliki lima tingkatan yaitu, TK, SD, SMP, SMA dan Perguruan Tinggi. Perguruan tinggi yang ada, negeri maupun swasta, di dalam maupun di luar negeri jumlahnya sangat banyak dan memiliki keunggulan serta kelebihan masing-masing. Seperti dunia bisnis dunia pendidikan juga saling bersaing, sehingga masyarakat dapat memilih Perguruan Tinggi mana yang lebih baik dan yang sesuai keinginannya. Hal ini yang menyebabkan banyaknya jumlah calon mahasiswa baru yang mendaftar di banyak Perguruan Tinggi. Satu calon mahasiswa baru dapat mendaftar lebih dari satu Perguruan Tinggi. Jika lulus di lebih dari satu Perguruan Tinggi kemungkinan mereka akan menjalani semuanya atau melakukan pengunduran diri sebagai calon mahasiswa di salah satu Perguruan Tinggi.

Hal ini sangat berdampak pada Perguruan Tinggi tersebut jika jumlah maksimal calon mahasiswa baru telah ditetapkan, dan pada saat pendaftaran jumlahnya lebih dari yang diperkirakan tetapi pada saat registrasi banyak calon mahasiswa baru yang mengundurkan diri dengan cara tidak melakukan registrasi. Hal ini juga berarti secara potensial Perguruan Tinggi ini tidak dapat mempertahankan calon mahasiswa baru tersebut. Jika kemungkinan pengunduran diri seorang calon mahasiswa baru dapat diketahui lebih dini dan Perguruan Tinggi dapat memperoleh informasi mengenai jumlah calon mahasiswa baru yang akan melakukan registrasi secepat mungkin dari daftar calon mahasiswa yang dinyatakan lulus dari seleksi penerimaan mahasiswa baru, maka pihak manajemen Perguruan Tinggi dapat melakukan tindakan-tindakan yang diperlukan untuk mempertahankan calon-calon mahasiswa tersebut.

Salah satu cara untuk melakukan analisis kemungkinan pengunduran diri seorang calon mahasiswa baru di suatu perguruan tinggi maka dibutuhkan suatu metode sebagai alat bantu dalam menggali berbagai informasi dari data tersebut. Alat bantu yang dapat digunakan

adalah metode decision tree untuk *data mining* menggunakan algoritma C4.5. *Decision tree* merupakan teknik klasifikasi penting dalam klasifikasi data mining. *Decision tree* telah terbukti menjadi alat yang berharga untuk klasifikasi, deskripsi, dan generalisasi data. *Decision tree* untuk kumpulan data terdapat pada banyak disiplin ilmu seperti pemrosesan sinyal, pengenalan pola, teori keputusan, statistik, pembelajaran mesin dan jaringan syaraf tiruan. (Karim dan Duaimi, 2014).

Data mining adalah salah satu bidang ilmu untuk menganalisa data dari perspektif berbeda dan menyimpulkannya menjadi informasi-informasi yang penting yang dapat dipakai untuk meningkatkan keuntungan, memperkecil biaya pengeluaran atau keduanya (Angga dan Riani, 2012). Data mining telah berhasil diterapkan pada banyak bidang, keseluruhan tujuan proses data mining adalah mengekstrak informasi dari kumpulan data dan mengubahnya menjadi struktur yang mudah dimengerti untuk digunakan lebih jauh. Data mining terutama digunakan untuk model klasifikasi dan prediksi. Klasifikasi adalah suatu bentuk analisis data yang mengekstrak model yang menggambarkan kelas data penting. C4.5 adalah salah satu algoritma klasifikasi yang paling klasik pada data mining (Agrawal dan Gupta, 2013).

Penerapan *data mining* telah dilakukan oleh Wahyudi, dkk (2011) yaitu untuk mengetahui program studi yang menjadi banyak peminat di masa yang akan datang bagi calon mahasiswa baru di suatu Perguruan Tinggi. Dalam kasus ini metode *data mining* yang digunakan adalah *klasifikasi*. Dimana prosesnya melakukan klasifikasi dari kumpulan data calon mahasiswa yang ada. Model klasifikasi yang digunakan adalah pendekatan pohon keputusan (*decision tree*). Algoritma yang dipakai sebagai algoritma pembentuk pohon keputusannya adalah algoritma C4.5.

Penerapan algoritma C4.5 juga pernah dilakukan oleh Nurma Jayanti, Sulisty Puspidjati dan Tety Elida (2008) dalam memprediksi kebangkrutan bank berdasarkan rasio keuangan bank. Dengan bantuan perangkat lunak, dan algoritma C.45 pada teknik *data mining* yang berbasis *decision tree* maka dapat dilakukan analisa kemungkinan pengunduran diri seorang calon mahasiswa baru dapat diketahui lebih dini di suatu Perguruan Tinggi sehingga pihak manajemen Perguruan Tinggi dapat melakukan tindakan-tindakan untuk mempertahankan calon mahasiswa tersebut di masa yang akan datang.

Pada penelitian ini diusulkan Sistem Prediksi Pengunduran Diri Calon Mahasiswa Baru Menggunakan Algoritma C45. Sistem ini di buat dengan menggunakan algoritma C45 untuk mendapatkan model rule yang digunakan untuk menentukan kelas dari data baru. Dengan adanya sistem ini diharapkan dapat memprediksi calon mahasiswa baru di Politeknik Negeri Bengkalis, dengan hasil apakah calon mahasiswa tersebut melakukan Daftar ulang atau tidak.

2. TINJAUAN PUSTAKA

Penelitian dibawah merupakan penelitian-penelitian yang telah dilakukan sebelumnya yang berkaitan dengan permasalahan pada penelitian yang akan dilakukan. Salah satu cara untuk melakukan analisis kemungkinan pengunduran diri seorang calon mahasiswa baru di suatu perguruan tinggi maka dibutuhkan suatu metode sebagai alat bantu dalam menggali berbagai informasi dari data tersebut. Alat bantu yang dapat digunakan adalah metode decision tree untuk data mining menggunakan algoritma C4.5. *Decision tree* merupakan teknik klasifikasi penting dalam klasifikasi data mining. *Decision tree* telah terbukti menjadi alat yang berharga untuk klasifikasi, deskripsi, dan generalisasi data. *Decision tree* untuk kumpulan data terdapat pada banyak disiplin ilmu seperti pemrosesan sinyal, pengenalan pola, teori keputusan, statistik, pembelajaran mesin dan jaringan syaraf tiruan. (Karim dan Duaimi, 2014).

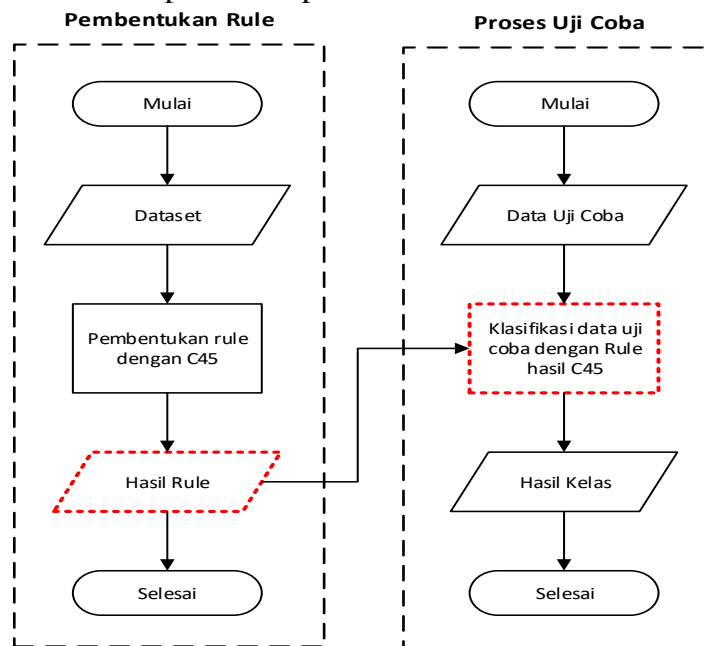
Data mining adalah proses menganalisa data dari perspektif berbeda dan menyimpulkannya menjadi informasi-informasi yang penting yang dapat dipakai untuk meningkatkan keuntungan, memperkecil biaya pengeluaran atau keduanya (Angga dan Riani, 2012). Data mining telah berhasil diterapkan pada banyak bidang, keseluruhan tujuan proses

data mining adalah mengekstrak informasi dari kumpulan data dan mengubahnya menjadi struktur yang mudah dimengerti untuk digunakan lebih jauh. Data mining terutama digunakan untuk model klasifikasi dan prediksi. Klasifikasi adalah suatu bentuk analisis data yang mengekstrak model yang menggambarkan kelas data penting. C4.5 adalah salah satu algoritma klasifikasi yang paling klasik pada data mining (Agrawal dan Gupta, 2013).

Penerapan data mining telah dilakukan oleh Wahyudi, dkk (2011) yaitu untuk mengetahui program studi yang menjadi banyak peminat di masa yang akan datang bagi calon mahasiswa baru di suatu Perguruan Tinggi. Dalam kasus ini metode data mining yang digunakan adalah klasifikasi. Dimana prosesnya melakukan klasifikasi dari kumpulan data calon mahasiswa yang ada. Model klasifikasi yang digunakan adalah pendekatan pohon keputusan (decision tree). Algoritma yang dipakai sebagai algoritma pembentuk pohon keputusannya adalah algoritma C4.5. Penerapan algoritma C4.5 juga pernah dilakukan oleh Nurma Jayanti, Sulisty Puspitodjati dan Tety Elida (2008) dalam memprediksi kebangkrutan bank berdasarkan rasio keuangan bank. Berdasarkan analisa pada penelitian sebelumnya maka algoritma C45 dapat diterapkan untuk memprediksi calon mahasiswa baru di Politeknik Negeri Bengkalis, apakah calon mahasiswa tersebut melakukan daftar ulang atau tidak.

3. METODE PENELITIAN

Pada tahap ini akan dijelaskan tahapan metode penelitian yang diusulkan. Pada penelitian ini ada 2 tahapan yaitu pembentukan rule dengan menggunakan metode C45 dan proses uji coba data tes menggunakan rule hasil dari C45. Adapun kedua tahapan yang diusulkan pada penelitian ini dapat dilihat pada Gambar 1.



Gambar 1 Diagram sistem yang diusulkan.

A. Pembentukan Rule

Pembentukan rule adalah tahap yang bertujuan untuk membuat sebuah model rule yang akan digunakan untuk model rule sebagai klasifikasi terhadap data tes.

1. Dataset

Dataset yang digunakan untuk membentuk rule adalah data calon mahasiswa yang lulus seleksi di Politeknik Negeri Bengkalis. Data yang digunakan adalah data tahun 2016 dan 2017.

2. Pembentukan Rule dengan C45

Pembentukan rule dilakukan dengan menggunakan tools WEKA. Adapun metode yang digunakan untuk membuat rule adalah C45 atau dikenal dengan J48 di WEKA dengan menggunakan *cross validation* 10. Variabel-variabel yang digunakan adalah jenis kelamin,

kota asal, jurusan di SMU, pilihan kampus 1, pilihan prodi 1, pilihan kampus 2, pilihan prodi 2 dan lulus diprodi.

3. Hasil Rule

Hasil rule merupakan model rule yang didapatkan dari analisa pembentukan rule dengan menggunakan metode C45. Hasil rule dapat ditunjukkan pada Tabel 1.

Tabel 1 Model rule yang digunakan

KOTA = DUMAI: DAFTAR (27.0/10.0)	KOTA = PADANG LAWAS: DAFTAR (1.0)
KOTA = BENGKALIS: DAFTAR (298.0/89.0)	KOTA = ROKAN HILIR: TIDAK (10.0/4.0)
KOTA = KUANTAN SINGINGI: DAFTAR (1.0)	KOTA = TAPANULI UTARA: DAFTAR (2.0/1.0)
KOTA = PASAMAN BARAT: TIDAK (21.0/7.0)	KOTA = MANDAILING NATAL: TIDAK (15.0/2.0)
KOTA = KEPULAUAN MERANTI	KOTA = PADANG: TIDAK (4.0/1.0)
JURUSAN ASAL = IPS: DAFTAR (4.0/2.0)	KOTA = LIMA PULUH KOTA: TIDAK (4.0/1.0)
JURUSAN ASAL = IPA: DAFTAR (8.0/1.0)	KOTA = DELI SERDANG: DAFTAR (1.0)
JURUSAN ASAL = ADMINISTRASI DAN PERKANTORAN: DAFTAR (1.0)	KOTA = DHARMASRAYA: DAFTAR (1.0)
JURUSAN ASAL = AKUNTANSI: TIDAK (1.0)	KOTA = SOLOK SELATAN: TIDAK (7.0/1.0)
JURUSAN ASAL = GEOLOGI PERTAMBANGAN: DAFTAR (0.0)	KOTA = -: TIDAK (9.0/3.0)
JURUSAN ASAL = TEKNIK INSTALASI TENAGA LISTRIK: DAFTAR (6.0)	KOTA = PEKAN BARU: TIDAK (3.0/1.0)
JURUSAN ASAL = TEKNIK AUDIO-VIDEO: DAFTAR (0.0)	KOTA = MEDAN: DAFTAR (1.0)
JURUSAN ASAL = TEKNIK SURVEY DAN PEMETAAN: DAFTAR (0.0)	KOTA = TAPANULI TENGAH: DAFTAR (1.0)
JURUSAN ASAL = TEKNIK KOMPUTER DAN JARINGAN: TIDAK (4.0)	KOTA = LABUHAN BATU: TIDAK (2.0)
JURUSAN ASAL = TEKNIK PRODUKSI PERMINYAKAN: DAFTAR (0.0)	KOTA = PASAMAN: TIDAK (12.0)
JURUSAN ASAL = TEKNIK PEMESINAN: DAFTAR (0.0)	KOTA = SIJUNJUNG: TIDAK (2.0)
JURUSAN ASAL = TEKNIK PEMBORAN MINYAK: DAFTAR (0.0)	KOTA = ASAHAN: TIDAK (1.0)
JURUSAN ASAL = TEKNIK SEPEDA MOTOR: DAFTAR (0.0)	KOTA = HUMBANG HASUNDUTAN: TIDAK (2.0)
JURUSAN ASAL = TEKNIK PENGELASAN: DAFTAR (0.0)	KOTA = BUKITTINGGI: TIDAK (1.0)
JURUSAN ASAL = AGRIBISNIS DAN HOLTIKULTURA: DAFTAR (0.0)	KOTA = SOLOK: TIDAK (1.0)
JURUSAN ASAL = 0.0: DAFTAR (0.0)	KOTA = PAYAKUMBUH: TIDAK (1.0)
JURUSAN ASAL = -: DAFTAR (0.0)	KOTA = BOGOR: TIDAK (1.0)
JURUSAN ASAL = TEKNIK GAMBAR BANGUNAN: DAFTAR (0.0)	KOTA = PESISIR SELATAN: TIDAK (3.0)
JURUSAN ASAL = NAUTIKA KAPAL NIAGA: DAFTAR (0.0)	KOTA = OGAN KOMERING ILIR: TIDAK (1.0)
JURUSAN ASAL = AKOMODASI PERHOTELAN: DAFTAR (0.0)	KOTA = AGAM: TIDAK (1.0)
JURUSAN ASAL = TEKNIK ELEKTRONIKA INDUSTRI: DAFTAR (0.0)	KOTA = SIMALUNGUN: TIDAK (1.0)
JURUSAN ASAL = TEKNIK KENDARAAN RINGAN: TIDAK (2.0)	KOTA = B A T A M: TIDAK (1.0)
JURUSAN ASAL = REKAYASA PERANGKAT LUNAK: DAFTAR (0.0)	KOTA = PARIAMAN: TIDAK (1.0)
JURUSAN ASAL = TEKNIK OTOMASI INDUSTRI: DAFTAR (0.0)	KOTA = SERDANG BEDAGAI: TIDAK (1.0)
KOTA = TANAH DATAR: TIDAK (3.0/1.0)	KOTA = TEBING TINGGI: TIDAK (1.0)
KOTA = ENREKANG: DAFTAR (3.0)	KOTA = BEKASI: TIDAK (1.0)
KOTA = PADANG PARIAMAN: TIDAK (5.0/1.0)	KOTA = PADANG PANJANG: TIDAK (1.0)
KOTA = SIAK: DAFTAR (19.0/8.0)	

B. Proses Uji Coba

1. Data Uji Coba

Data uji coba merupakan data calon mahasiswa yang akan di klasifikasi. Adapun data uji yang digunakan berupa data jenis kelamin, kota asal, jurusan di SMU, pilihan kampus 1, pilihan prodi 1, pilihan kampus 2, pilihan prodi 2, dan lulus di prodi.

2. Klasifikasi dengan Rule

Data uji coba yang diinputkan akan diklasifikasi menggunakan rule yang telah ditentukan. Tujuan dari klasifikasi ini adalah untuk menentukan kelas dari data uji coba.

3. Hasil Kelas

Kelas yang dihasilkan merupakan hasil klasifikasi terhadap data uji coba. Adapun kelas yang dihasilkan adalah DAFTAR ULANG dan TIDAK DAFTAR ULANG.

4. HASIL UJI COBA DAN PEMBAHASAN

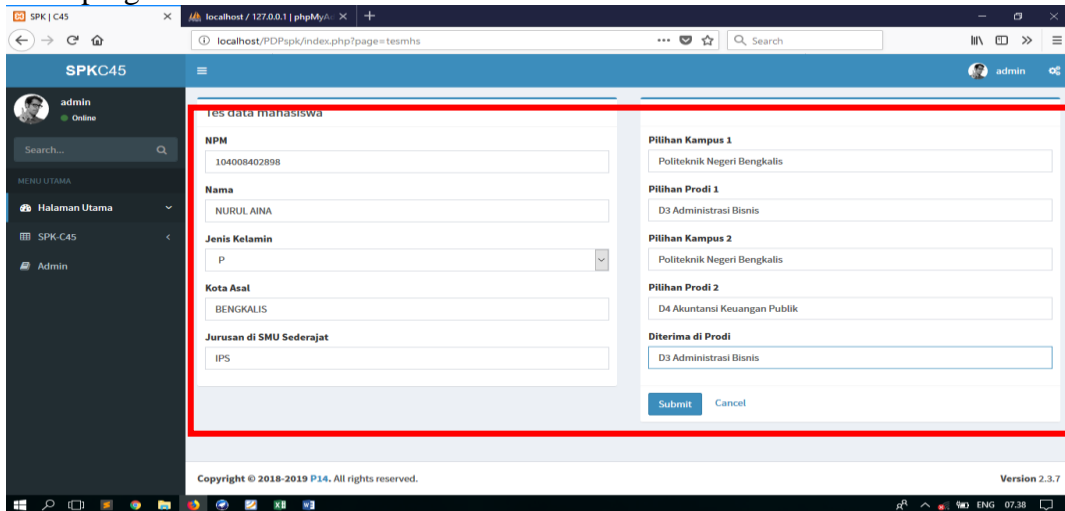
Pada tahap ini akan dibahas tentang implementasi hasil uji coba dan pembahasannya.

A. Implementasi

Uji coba dilakukan dengan dua skenario yaitu skenario uji coba per data dan skenario uji coba banyak data. Uji coba per data merupakan uji coba yang dilakukan terhadap data satu per satu, sedangkan uji coba banyak data merupakan uji coba yang dilakukan banyak data uji dalam sekali uji coba.

1. UjiCoba per Data

Pada Gambar 2 menunjukkan sebuah implementasi dari proses input data uji coba. Data uji coba yang digunakan merupakan data calon mahasiswa. Data calon mahasiswa yang diinputkan berupa NPM, Nama mahasiswa, Jenis kelamin, Kota asal, Jurusan waktu SMU, Pilihan kampus pertama, Pilihan program studi pertama, Pilihan kampus kedua, Pilihan program studi kedua, dan program studi dimana calon mahasiswa tersebut lulus seleksi atau diterima di program studi.



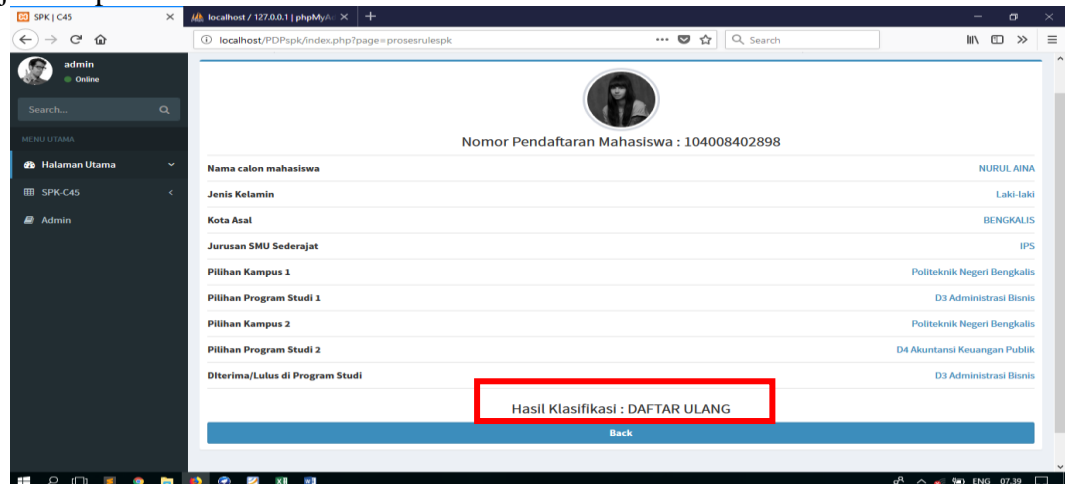
The screenshot shows a web browser window with the URL localhost/127.0.0.1/phpMyA. The page title is 'SPKC45'. The user is logged in as 'admin'. The main content area is titled 'Tes data mahasiswa' and contains a form with the following fields:

NPM	104008402898
Nama	NURULAINA
Jenis Kelamin	P
Kota Asal	BENGKALIS
Jurusan di SMU Sederajat	IPS
Pilihan Kampus 1	Politeknik Negeri Bengkalis
Pilihan Prodi 1	D3 Administrasi Bisnis
Pilihan Kampus 2	Politeknik Negeri Bengkalis
Pilihan Prodi 2	D4 Akuntansi Keuangan Publik
Diterima di Prodi	D3 Administrasi Bisnis

Buttons for 'Submit' and 'Cancel' are located at the bottom right of the form.

Gambar 2 Input data calon mahasiswa / data uji coba.

Proses klasifikasi merupakan proses penentuan kelas terhadap data ujicoba yang diinputkan. Terdapat dua kelas pada klasifikasi ini yaitu kelas DAFTAR ULANG dan kelas TIDAK DAFTAR ULANG. Adapun hasil proses klasifikasi yang dilakukan dapat ditunjukkan pada Gambar 3.



The screenshot shows a web browser window with the URL localhost/PDPspk/index.php?page=proserulespk. The page displays the student's details and the classification result:

Nomor Pendaftaran Mahasiswa : 104008402898

Nama calon mahasiswa	NURULAINA
Jenis Kelamin	Laki-laki
Kota Asal	BENGKALIS
Jurusan SMU Sederajat	IPS
Pilihan Kampus 1	Politeknik Negeri Bengkalis
Pilihan Program Studi 1	D3 Administrasi Bisnis
Pilihan Kampus 2	Politeknik Negeri Bengkalis
Pilihan Program Studi 2	D4 Akuntansi Keuangan Publik
Diterima/Lulus di Program Studi	D3 Administrasi Bisnis

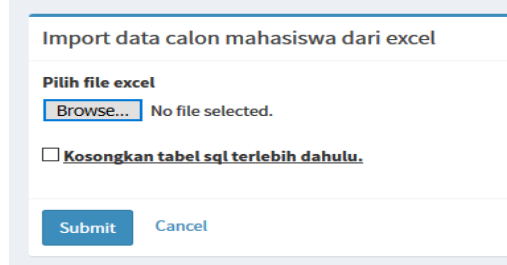
The classification result is: Hasil Klasifikasi : DAFTAR ULANG

A 'Back' button is located at the bottom of the page.

Gambar 3 Hasil klasifikasi.

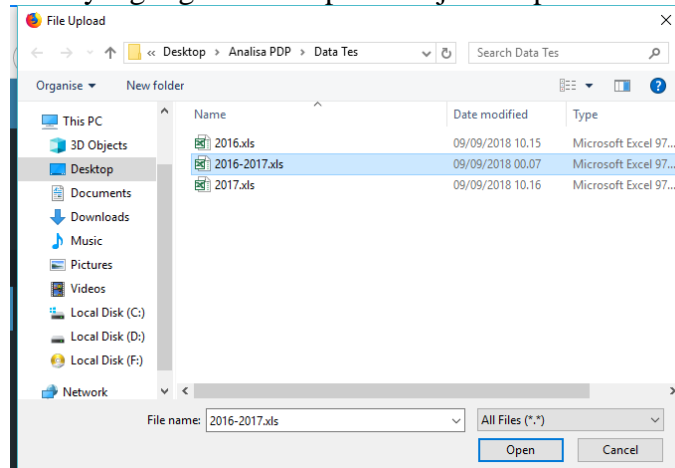
2. UjiCobaBanyak Data

Uji coba banyak data merupakan uji coba yang dilakukan dengan cara menguji coba banyak data dalam sekali uji coba. Uji coba ini dilakukan dengan cara meng-import data dari excel ke database. Kemudian data yang ada dalam database akan di klasifikasi untuk menentukan kelas dari data-data yang ada didalam database tersebut. Adapun form import data dari data excel dapat ditunjukkan pada Gambar 4.



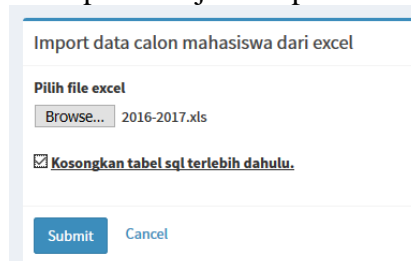
Gambar 4 Form import data uji coba dari excel.

Data uji coba yang digunakan merupakan data calon mahasiswa tahun 2016, 2017, dan 2016&2017. Adapun data yang digunakan dapat ditunjukkan pada Gambar 5.



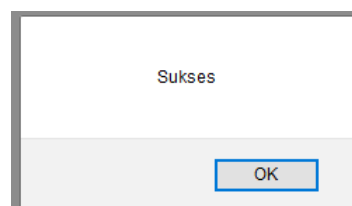
Gambar 5 Data uji coba.

Sebelum import data maka data yang ada dalam database dihapus terlebih dahulu. Hal ini dilakukan agar data uji coba yang baru tidak bercampur dengan data yang lama. Adapun proses import data dan hapus data dapat ditunjukkan pada Gambar 6.



Gambar 6 Proses import data uji coba.

Proses import data uji coba membutuhkan waktu beberapa detik dan tergantung dengan banyaknya data yang di import. Setelah berhasil import data maka akan muncul pesan seperti ditunjukkan pada Gambar 7.



Gambar 7 Import data uji coba berhasil.

Setelah proses import data uji coba berhasil maka dapat ditampilkan di sistem seperti ditunjukkan pada Gambar 8. Data yang sudah ada di dalam database dapat dilakukan uji coba dengan melakukan klasifikasi semua data yang ada didalam database untuk menentukan kelas dari data-data tersebut. Proses klasifikasi dapat dilakukan dengan cara klik tombol Klasifikasi seperti ditunjukkan pada Gambar 8.

Data Calon Mahasiswa Baru

Buttons: + Add, - Delete, Import from xls, **Klasifikasi**

Show: 10 entries

Aksi	No	NPM	Nama	Jenis Kelamin	Diterima di Prodi	Status DU
	1	698679345493	Wella Pebriana	L	D4-Teknik Listrik	TIDAK
	2	698573032037	Khoirul Hakim	L	D4-Rekayasa Perangkat Lunak	DAFTAR
	3	698573031628	Nur Asikin	P	D4-Administrasi Bisnis Internasional	DAFTAR
	4	698315092146	DESTI KRISDA RAMADANI	P	D4 Teknik Mesin Produksi & Perawatan	TIDAK
	5	698227064230	MIRA MILARTI	P	D3 Administrasi Bisnis	TIDAK
	6	697869692767	Sri Rahayuni	P	D4-Akuntansi Keuangan Publiik	DAFTAR
	7	697746436710	Nursehati	P	D3-Teknik Informatika	TIDAK
	8	697746431387	Ratna Sahara	P	D3-Administrasi Bisnis	TIDAK
	9	697735669922	Winda Sari	P	D3-Administrasi Bisnis	DAFTAR
	10	697735669572	Arie Isnawan	L	D3-Teknik Mesin	DAFTAR

Showing 1 to 10 of 988 entries

Gambar 8 Data uji cobahasil import.

Hasil proses klasifikasi data uji coba dapat dilihat langsung ketika proses klasifikasi selesai. Adapun hasil dari klasifikasi dapat ditunjukkan pada Gambar 9.

Proses Klasifikasi C45 Data calon MHS

Hasil Proses Klasifikasi Rule C45

No	NPM	Nama	Program Studi	Status DU	Hasil Klasifikasi	Status Klasifikasi
1	102027659376	SAFIRANNUR	D3 Bahasa Inggris Bisnis	TIDAK	TIDAK	Benar
2	102040539450	Isnaida Marisa	D3-Administrasi Bisnis	TIDAK	TIDAK	Benar
3	102053782354	SURIYANTO	D3 Teknik Perkapalan	TIDAK	TIDAK	Benar
4	102053808852	DAMERIA SANTY SIRAIT	D3 Administrasi Bisnis	DAFTAR	TIDAK	Salah
5	102053809012	YOLANDA THERESIA PARAPAT	D3 Teknik Sipil	TIDAK	TIDAK	Benar
6	102053809325	Monaria Br Simaremare	D3-Teknik Sipil	TIDAK	TIDAK	Benar
7	102053855775	Aida Fitri Nasution	D3-Administrasi Bisnis	TIDAK	TIDAK	Benar
8	102053861325	AGUS SURİYANTO	D3 Nautika	TIDAK	TIDAK	Benar
9	102053863799	ANDRI ANNURU	D3 Teknik Perkapalan	TIDAK	TIDAK	Benar
10	102053864905	ARI KUSWARA	D4 Teknik Mesin Produksi & Perawatan	DAFTAR	TIDAK	Salah

Showing 1 to 10 of 988 entries

Gambar 9 Hasil klasifikasi.

Pada Gambar 9 menunjukkan bahwa kolom Status DU merupakan kelas sebenarnya dari data uji coba. Kolom Hasil klasifikasi merupakan kelas dari hasil klasifikasi yang dilakukan terhadap data uji coba. Kolom Status klasifikasi merupakan hasil penentuan kelas dari proses klasifikasi. Untuk menentukan hasil klasifikasi benar maka kelas hasil klasifikasi dibandingkan kelas sebenarnya, jika keduanya memiliki hasil yang sama maka proses klasifikasi dianggap benar, dan jika keduanya tidak sama maka proses klasifikasi dianggap salah.

Untuk mengukur keberhasilan sistem yang dibuat maka diukur dengan menggunakan akurasi. Adapun persamaan akurasi yang digunakan dapat dilihat pada persamaan berikut.

$$Akurasi = \frac{Total\ uji\ coba\ benar}{Total\ data\ uji\ coba} \times 100\%$$

Dimana :

Total uji coba benar adalah jumlah total dari data uji coba yang terklasifikasi dengan benar.

Total data uji coba adalah jumlah total dari data uji coba yang diklasifikasi.

Sedangkan untuk mengukur Error dapat ditunjukkan pada persamaan berikut.

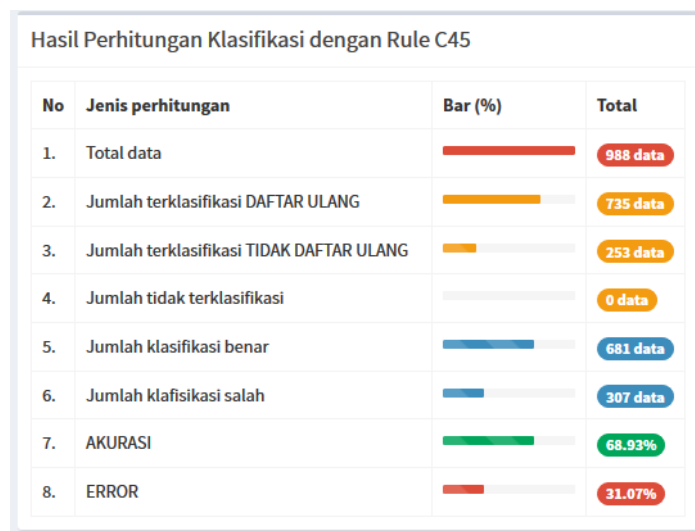
$$Error = \frac{\text{Total uji coba salah}}{\text{Total data uji coba}} \times 100\%$$

Dimana :

Total uji coba salah adalah jumlah total dari data uji coba yang tidak terklasifikasi dengan benar.

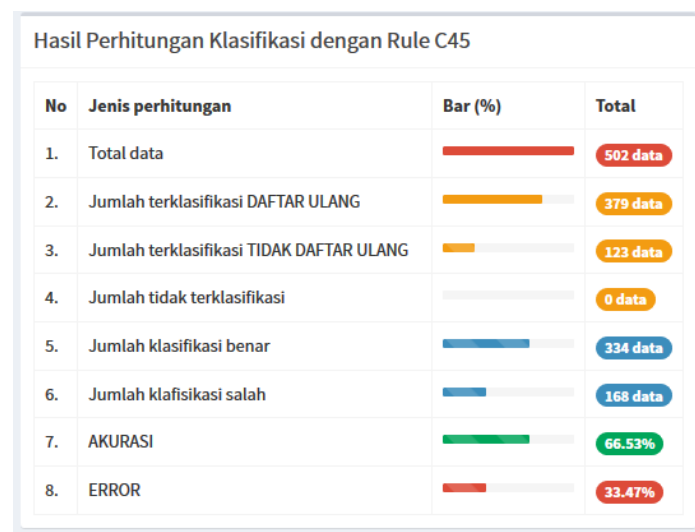
Total data uji coba adalah jumlah total dari data uji coba yang diklasifikasi.

Hasil akurasi uji coba terhadap data calon mahasiswa tahun 2016 & 2017 dapat ditunjukkan pada Gambar 10.



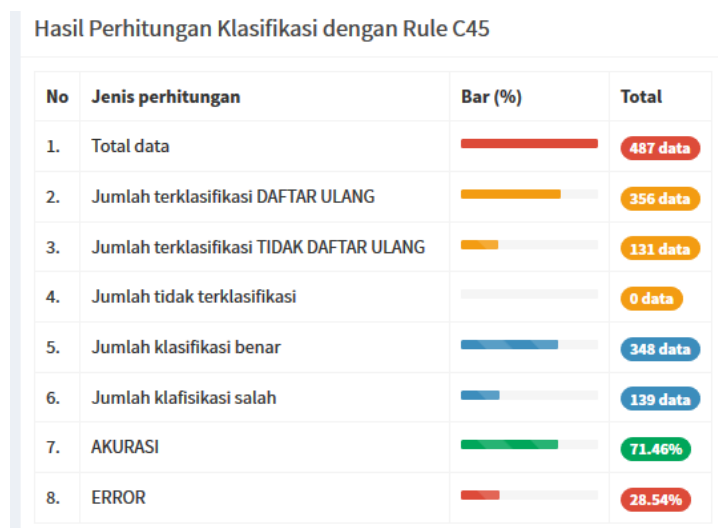
Gambar 10 Hasil perhitungan klasifikasi data 2016&2017.

Hasil akurasi uji coba terhadap data calon mahasiswa tahun 2016 dapat ditunjukkan pada Gambar 11.



Gambar 11 Hasil perhitungan klasifikasi data 2016.

Hasil akurasi uji coba terhadap data calon mahasiswa tahun 2016 & 2017 dapat ditunjukkan pada Gambar 12.



Gambar 12 Hasil perhitungan klasifikasi data 2017.

B. Pembahasan

Sistem yang diusulkan merupakan sistem pengambilan keputusan berbasis *decision tree* untuk pengunduran diri calon mahasiswa baru. Sistem ini dibangun menggunakan model rule hasil dari algoritma C45. Pembentukan model rule dilakukan dengan menganalisa variabel-variabel dari data calon mahasiswa tahun 2016 dan 2017 yang digunakan untuk mendapatkan model rule terbaik. Algoritma C45 juga dibandingkan dengan algoritma CART ketika analisa pembentukan rule, sehingga algoritma C45 memiliki akurasi lebih baik ketika dibandingkan dengan algoritma CART. Model rule terbaik kemudian diimplementasikan pada sebuah sistem yang berbasis WEBSITE.

Ada beberapa tahap uji coba yang dilakukan terhadap sistem yang diusulkan. Adapun uji coba dilakukan terhadap data uji coba tahun gabungan (2016&2017), data uji coba tahun 2016, dan data uji coba tahun 2107. Dari hasil uji coba yang telah dilakukan didapatkan hasil perhitungan akurasi dan error dari proses klasifikasi. Adapun hasil perbandingan uji coba yang telah dilakukan dapat ditunjukkan pada Tabel 2.

Tabel 2 Hasil perbandingan uji coba.

No	Data (tahun)	Total data	Terklasifikasi BENAR	Terklasifikasi SALAH	Akurasi (%)	Error (%)
1	2016 - 2017	988	681	307	68,93	31,07
2	2016	502	334	168	66,53	33,47
3	2017	487	348	139	71,46	28,54

Dari hasil analisa yang telah dilakukan menunjukkan bahwa akurasi klasifikasi memiliki akurasi 71,46% dan tidak terlalu signifikan. Hal ini terjadi karena pola data yang tidak beraturan. Ada data yang sama tetapi tidak memiliki kelas yang sama. Contohnya banyak data mahasiswa yang sama-sama berasal dari bengkalis dan memiliki data yang similar atau mirip tetapi berbeda kelas, misalkan ada yang Daftar Ulang dan ada juga yang Tidak Daftar Ulang. Kemiripan data yang tinggi dan kelas yang berbeda sangat mempengaruhi akurasi penentuan kelas dari data uji coba.

5. KESIMPULAN DAN SARAN

A. Kesimpulan

Penelitian ini bertujuan untuk menganalisa variabel-variabel dan metode decision tree apa yang memiliki akurasi klasifikasi yang tinggi pada sistem pengambilan keputusan untuk menentukan calon mahasiswa baru daftar ulang atau tidak. Uji coba dilakukan sebanyak 5 model scenario uji coba, sehingga setiap skenario uji coba memiliki variabel-variabel yang berbeda-beda. Dari 5 skenario uji coba yang telah dilakukan menunjukkan bahwa kumpulan variabel Jenis kelamin, kota asal, jurusan asal di SMU sederajat, pilihan kampus 1, pilihan prodi 1, pilihan kampus 2, pilihan prodi 2 dan diterima di prodi memiliki akurasi yang lebih

baik dibandingkan dengan skenario uji coba yang lain. Penggunaan metode C45 memiliki akurasi paling tinggi dibandingkan dengan metode CART yaitu tingkat akurasi 64,4%. Penelitian ini menghasilkan model rule C45 yang diimplementasikan pada sebuah aplikasi sistem pengambilan keputusan calon mahasiswa daftar ulang atau tidak berbasis WEB. Uji coba terhadap model rule dilakukan dengan menggunakan data calon mahasiswa tahun 2016 dan 2017. Berdasarkan hasil uji coba yang telah dilakukan terhadap sistem pada data gabungan tahun 2016&2017 memiliki akurasi 68,93%, data tahun 2016 memiliki akurasi 66,53% dan data tahun 2017 memiliki akurasi 71,46%.

B. Saran

Untuk saran pengembangan dari analisa yang diusulkan adalah lakukan analisa menggunakan variabel-variabel yang belum di uji coba pada analisa ini. Adapun variabel yang belum seperti nilai prestasi pada saat sekolah SMU sederajat, karena nilai memiliki kemungkinan calon mahasiswa daftar ulang atau tidak. Menggunakan data calon mahasiswa yang lebih banyak tahun 2018 dst. Kemudian lakukan uji coba dengan menggunakan beberapa metode klasifikasi lainnya seperti K-nearest neighbor, Neural Network dan metode lainnya yang berhubungan dengan klasifikasi.

6. REFERENSI

- Karim. I. A dan Duaimi. M. G, 2014, *Improved Accuracy For Decision Tree Algorithm Based On Unsupervised Discretization*, *IJCSMC*, Vol. 3, Issue. 6, June 2014, pg.176 – 183.
- Agrawal.G.L, Gupta. H, 2013, *Optimization of C4.5 Decision Tree Algorithm for Data Mining Application*, *International Journal of Emerging Technology and Advanced Engineering Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 3, Issue 3, March 2013)*.
- Angga Ginanjar, Riani, dkk. 2012. “*Penerapan Data Mining Untuk Memprediksi Kriteria Nasabah Kredit*” *Jurnal Universitas Komputer Indonesia*: Bandung.
- Wahyudi, Eko Nur, dkk. 2011. “*Analisa Profil Data Mahasiswa Baru terhadap Program Studi yang dipilih di Perguruan Tinggi Swasta Jawa Tengah dengan Menggunakan Teknik Data Mining*”. *Jurnal Universitas Stikubank*: Semarang.
- Jayanti, Nurma, dkk. 2008. “*Teknik Klasifikasi Pohon Keputusan untuk Memprediksi Kebangkrutan Bank Berdasarkan Rasio Keuangan Bank*”. *Jurnal Universitas Gunadarma*: Jakarta.