

Penerapan Algoritma C4.5 untuk Klasifikasi Jenis Pekerjaan Alumni di Universitas Muhammadiyah Yogyakarta

(Application of C4.5 Algorithm for Classification of Types of Alumni Work at Universitas Muhammadiyah Yogyakarta)

ASRONI, BADRAHINI MASAJENG RESPATI, SLAMET RIYADI

ABSTRACT

The development of education in Indonesia has increased very rapidly. One of the things that have become a benchmark for success in the quality of education at the university is the kind of job getting graduates after graduation. This research aims to identify factors that have an impact on the type of job classification method based on the C 4.5 alumni algorithm. The methodology of this research begins with the study of literature, the identification of a process of data extraction, data selection, data collection, data processing, data testing, and DA conclusion. This research uses some features of the data on a few faculty members, the year of graduation, the annual completion rate, and the strength as a classification performance parameter. Graduates data used up to 259, and consisted of 3 faculties of Economics, medicine and engineering forces from 2001-2013 and graduated from 2011-2016. The research results that have been done is if it comes from the Faculty of Economics, in 2011 and 2012 the majority of work in the private sector has passed, if it comes from the Faculty of Medicine with the years 2011 and 2012 graduated with a cumulative labor rate of between 3 to 3.5 majority working in The private sector, 2012 with a GPA between 3 and 3.5 working in the Private Sector. Finally, the C4.5 algorithm is suitable for the classification of alumni work types.

Keywords: Algoritma C4.5, Gain ratio, Pohon keputusan, Data mining, Alumni

PENDAHULUAN

Perkembangan dunia pendidikan di Indonesia mengalami peningkatan yang sangat pesat. Salah satu hal yang menjadi tolok ukur keberhasilan mutu pendidikan suatu Universitas adalah jenis pekerjaan alumni yang didapatkan setelah lulus. Tentunya hal tersebut juga terjadi di Universitas Muhammadiyah Yogyakarta. Banyaknya jumlah alumni di UMY memiliki peran penting terhadap peningkatan mutu pendidikan. Hal yang dapat dijadikan sebagai tolok ukurnya adalah jenis pekerjaan yang diperoleh alumni setelah lulus, tetapi masalahnya belum dilakukan pengolahan pada data alumni dengan menggunakan metode data mining.

Berdasarkan uraian diatas, klasifikasi jenis pekerjaan alumni belum pernah dilakukan sebelumnya, oleh karena itu penulis akan

menggunakan algoritma C4.5 dalam melakukan klasifikasi jenis pekerjaan alumni dengan hasil berupa pohon keputusan yang dapat digunakan untuk mengidentifikasi dan melihat hubungan antara faktor-faktor yang mempengaruhi suatu kasus (Andriani 2012).

Tujuan dari penelitian ini adalah mengimplementasikan metode klasifikasi menggunakan algoritma C4.5 dalam melakukan klasifikasi data jenis pekerjaan alumni fakultas Ekonomi, fakultas Teknik, dan fakultas Kedokteran di UMY. Dan mengetahui faktor-faktor yang berpengaruh pada jenis pekerjaan alumni berdasarkan hasil pohon keputusan dari algoritma C4.5 pada klasifikasi data jenis pekerjaan alumni UMY.

Data mining

Data mining merupakan proses penggalian dan pertambangan pengetahuan dari sejumlah data yang besar, *database* atau *repository* database

lainnya (Santoso, n.d.). Tujuan utama dari penambahan data ini untuk menemukan pengetahuan baru yang tersembunyi dari *database* tersebut (Elisa 2017).

Klasifikasi

Klasifikasi adalah suatu teknik yang dilakukan untuk memprediksi *class* atau properti dari setiap *instace* data. Teknik ini dilakukan pada data baru dengan memanipulasi data yang ada dan telah diklasifikasikan menggunakan hasilnya untuk memberikan sejumlah aturan.

a. Konsep Entropy

Entropy (S) merupakan jumlah bit yang diperkirakan dibutuhkan untuk dapat mengekstrak suatu kelas (+ atau -) dari sejumlah data acak pada ruang sampel S. *Entropy* dapat dikatakan sebagai kebutuhan bit untuk menyatakan suatu kelas. Semakin kecil nilai *entropy* maka akan semakin baik *entropy* yang digunakan dalam mengekstrak suatu kelas. *Entropy* digunakan untuk mengukur ketidakkaslian S (Nugroho 2014). Perhitungan *Entropy* seperti pada Persamaan 1.

$$Entropy(S) = \sum_{i=0}^n -p_i * \log_2(p_i) \quad (1)$$

b. Information Gain

Information Gain adalah informasi yang didapatkan dari perubahan *entropy* di suatu kumpulan data, baik melalui observasi maupun disimpulkan dengan cara melakukan partisipasi terhadap suatu set data (Ginting, Zarman, and Hamidah 2014). Perhitungan *Information Gain* seperti pada Persamaan 2.

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S) \quad (2)$$

c. SplitInfo

SplitInfo merupakan rumus yang menyatakan informasi potensial atau *entropy*. dapat dilihat dalam rumus (3) (Widayu et al. 2017). Perhitungan *SplitInfo* seperti pada Persamaan 3.

$$SplitInfo(S, A) = -\sum_{i=1}^n \frac{S_i}{S} \log_2 \frac{S_i}{S} \quad (3)$$

d. Gain Ratio

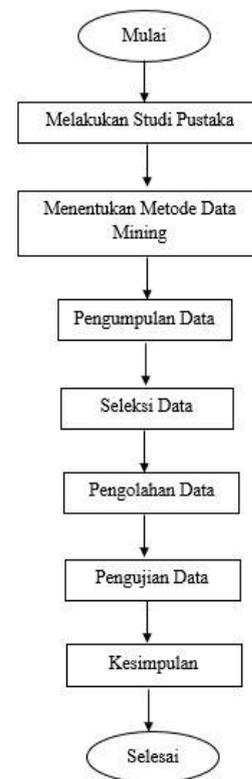
Gain Ratio adalah modifikasi dari *information gain* yang digunakan untuk mengurangi bias atribut yang memiliki banyak cabang. *Gain*

ratio memiliki sifat: bernilai besar jika data menyebar rata dan bernilai kecil jika semua data masuk ke dalam satu cabang (Mabrur 2012). Perhitungan *Gain Ratio* seperti pada Persamaan 4.

$$GainRatio(S, A) = \frac{Gain(S,A)}{SplitInfo(S,A)} \quad (4)$$

METODOLOGI PENELITIAN

Penerapan metodologi penelitian dilakukan dengan melakukan studi pustaka, menentukan metode data mining, pengumpulan data, seleksi data, pengolahan data, pengujian data, dan membuat kesimpulan seperti pada Gambar 1 (Fansuri 2012).



GAMBAR 1. Alur Metode Penelitian

1. Melakukan Studi Pustaka

Pertama hal yang dilakukan adalah melakukan studi pustaka yang berkaitan dengan penelitian-penelitian sebelumnya tentang penggunaan algoritma C4.5, serta untuk menguatkan penelitian ini berdasarkan teori yang digunakan. Dalam kasus ini peneliti ingin melakukan analisis pada klasifikasi jenis pekerjaan alumni di Universitas Muhammadiyah Yogyakarta dengan menggunakan algoritma C4.5 sekaligus

untuk mengetahui pola dan cara kerja klasifikasi dalam *data mining* tersebut (Kamagi and Hansun 2014).

2. Menentukan Metode Data Mining

Setelah melakukan studi pustaka tahap selanjutnya adalah menentukan metode yang sesuai untuk digunakan dalam teknik klasifikasi, setelah melakukan pengamatan dan observasi peneliti memilih menggunakan algoritma C4.5. Karena algoritma C4.5 dapat lakukan transformasi atau mengubah nilai atribut data ke dalam bentuk data yang sesuai agar data dapat diproses menggunakan algoritma C4.5. Sehingga akan diperoleh dataset utuh yang digunakan untuk proses ke tahapan selanjutnya.

3. Melakukan Tahapan Metode Klasifikasi

Berikut tahapan metode klasifikasi menggunakan algoritma C4.5 yang akan dilakukan peneliti untuk memperoleh hasil penelitian antara lain:

a. Melakukan Pengumpulan Data

Pada tahap ini peneliti melakukan pengumpulan data dan ini merupakan tahapan yang penting karena dapat berpengaruh terhadap hasil penelitian, sehingga dalam mengumpulkan data harus dilakukan dengan benar. Peneliti mendapatkan data ini dari *database server* BSI (Biro Sistem Informasi) Universitas Muhammadiyah Yogyakarta.

b. Melakukan Seleksi Data

Tahap ini dilakukan seleksi terhadap data *database excel*. Karena data yang diperoleh tidak semuanya digunakan, dipilih sesuai dengan atribut atau variabel yang dibutuhkan dalam penelitian dengan melakukan seleksi data sehingga menjadi *dataset*. Sebagai contoh pada penelitian ini atribut yang dipilih adalah fakultas, tahun lulus, ipk, angkatan, dan jenis pekerjaan. Pada tahap ini akan menghilangkan data yang *null*, data tidak valid, dan data yang ganda. karena data yang kosong ataupun tidak valid akan berpengaruh terhadap hasil yang diperoleh. Kemudian data akan diolah menggunakan *Microsoft Office Excel 2016*.

c. Melakukan Pengolahan Data

Setelah semua data yang diperlukan telah dipilih, maka tahap penelitian selanjutnya

adalah pengolahan data. Pada tahap ini akan dilakukan transformasi atau mengubah nilai atribut data ke dalam bentuk data yang sesuai agar data dapat diproses menggunakan algoritma C4.5. Sehingga akan diperoleh dataset utuh yang digunakan untuk proses ke tahapan selanjutnya.

d. Melakukan Pengujian Data

Pada tahap pengujian hasil akan dilakukan pengujian data baik secara manual dengan algoritma C4.5 dan menggunakan *software RapidMiner (Sembiring, Sibuea, and Sapta, n.d.)*.

e. Membuat Kesimpulan Penelitian

Berdasarkan hasil pengujian maka dapat ditarik kesimpulan yang mengacu pada rumusan masalah dan tujuan penelitian. Dan saran yang digunakan untuk mengembangkan penelitian selanjutnya serta dimasukkan untuk meningkatkan kualitas penelitian.

HASIL DAN PEMBAHASAN

Pengumpulan Data

Pengumpulan data dilakukan dengan cara meminta data ke Biro Sistem Informasi Universitas Muhammadiyah Yogyakarta. Data yang diperoleh adalah data alumni Universitas Muhammadiyah Yogyakarta dengan jumlah data awal 688 data alumni. Di dalam data alumni terdapat beberapa atribut antara lain: id, nim, nama, tempat lahir, jenis kelamin, id prodi, bulan lulus, tahun lulus, ipk, judul skripsi, alamat, email, nomor hp, nomor telepon, id jenis pekerjaan, nama institusi, alamat institusi, bagian, masa tunggu, id salary, tanggal lahir, tanggal mulai kerja, relevansi pendidikan, dan jabatan.

Penyeleksian Data

Pemilihan atribut atau variabel untuk penelitian ini dilakukan dengan pertimbangan atribut yang dipilih sudah sesuai dengan atribut yang dibutuhkan untuk melakukan proses klasifikasi jenis pekerjaan alumni. Dari data yang tersedia diambil beberapa atribut antara lain: id prodi, tahun lulus, ipk, angkatan, dan id jenis pekerjaan. Dari jumlah seluruh data yang ada, data yang digunakan jumlahnya 259 data alumni yang diambil dari 3 fakultas yaitu fakultas Ekonomi, fakultas Teknik, dan fakultas

Kedokteran dari angkatan 2001-2013 serta tahun lulus dari 2011-2016.

Pengolahan Data

Pada tahap ini dilakukan perubahan data atau transformasi data dengan mengubah nilai atribut yang awalnya bernilai nominal menjadi nilai-nilai atribut yang sesuai dengan data pada Tabel 1 dan data tidak ada *null*. Dan pada nilai atribut telah dirubah untuk memudahkan proses yang dilakukan dalam proses transformasi data.

Pengujian Data

Pada tahap pengujian data dilakukan dua pengujian yaitu menggunakan *software RapidMiner* dan menggunakan Algoritma C4.5, data pengujian seperti pada Tabel 1.

a. Tahapan Pra Proses

Pada Tabel 1 merupakan *dataset* yang digunakan untuk pra proses dari klasifikasi jenis pekerjaan alumni yang merupakan format data akhir yang didapat berdasarkan atribut yang telah dikelompokkan atau diklasifikasi.

b. Menghitung Entropy dan Gain Ratio

Tahap pembangunan pohon keputusan dilakukan setelah selesai menghitung nilai *entropy* total, nilai *entropy* dari tiap-tiap atribut, menghitung nilai *gain ratio*, menghitung nilai *split info*, dan menghitung nilai *gain ratio* serta selanjutnya

menentukan nilai *gain ratio* tertinggi, hal ini dapat dilihat di Tabel 2.

Dengan menggunakan rumus (1), maka nilai *Entropy* dapat dihitung. Maka menghitung nilai *Entropy* Total dilakukan dengan cara menghitung jumlah keputusan “PEMERINTAH” dan “SWASTA” dari seluruh jumlah kasus yang ada.

$$Entropy \text{ (Total)} = \left(\frac{-72}{259} \times \log_2 \left(\frac{72}{259} \right) \right) + \left(\frac{-187}{259} \times \log_2 \left(\frac{187}{259} \right) \right) = 0.85270$$

Hasil *Entropy* total diperoleh dari menghitung nilai total keputusan PEMERINTAH dengan jumlah 72 kasus dan SWASTA dengan jumlah 187 kasus, sedangkan 259 adalah jumlah dari keseluruhan kasus. Kemudian menghitung nilai *Entropy* dari masing-masing nilai atribut yang ada. Dengan menggunakan rumus (2) dapat menghitung nilai *Information Gain* tiap-tiap atribut.

$$Gain \text{ (Total, Fakultas)} = 0.85270 - \left(\left(\frac{99}{259} \times 0.68403 \right) + \left(\frac{63}{259} \times 0.63126 \right) + \left(\frac{97}{259} \times 0.99378 \right) \right) = 0.06549$$

$$Gain \text{ (Total, Tahun Lulus)} = 0.85270 - \left(\left(\frac{126}{259} \times 0.91829 \right) + \left(\frac{129}{259} \times 0.78244 \right) + \left(\frac{4}{259} \times 0 \right) \right) = 0.01625$$

$$Gain \text{ Ratio (Fakultas)} = \frac{0.06549}{1.55709} = 0.0426$$

$$Gain \text{ (Total, Angkatan)} = 0.85270 - \left(\left(\frac{16}{259} \times 0.68403 \right) + \left(\frac{243}{259} \times 0.86077 \right) \right) = 0.00208$$

TABEL 1. Data Pengujian

#	A	B	C	D	E
1	FAKULTAS	TAHUN LULUS	IPK	ANGKATAN	JENIS PEKERJAAN
2	Teknik	Lulusan 2011 dan 2012	IPK kurang dari 3	Angkatan 2001-2006	Swasta
3	Teknik	Lulusan 2011 dan 2012	IPK antara 3 sampai 3,5	Angkatan 2007-2013	Swasta
4	Kedokteran	Lulusan 2011 dan 2012	IPK antara 3 sampai 3,5	Angkatan 2007-2013	Swasta
5	Kedokteran	Lulusan 2011 dan 2012	IPK antara 3 sampai 3,5	Angkatan 2007-2013	Pemerintah
6	Kedokteran	Lulusan 2011 dan 2012	IPK antara 3 sampai 3,5	Angkatan 2007-2013	Swasta
7	Kedokteran	Lulusan 2011 dan 2012	IPK antara 3 sampai 3,5	Angkatan 2007-2013	Pemerintah
8	Kedokteran	Lulusan 2011 dan 2012	IPK kurang dari 3	Angkatan 2007-2013	Pemerintah
9	Kedokteran	Lulusan 2011 dan 2012	IPK antara 3 sampai 3,5	Angkatan 2007-2013	Pemerintah
10	Kedokteran	Lulusan 2011 dan 2012	IPK kurang dari 3	Angkatan 2007-2013	Swasta
11	Kedokteran	Lulusan 2011 dan 2012	IPK kurang dari 3	Angkatan 2007-2013	Swasta
12	Ekonomi	Lulusan 2011 dan 2012	IPK antara 3 sampai 3,5	Angkatan 2007-2013	Swasta
13	Teknik	Lulusan 2011 dan 2012	IPK antara 3 sampai 3,5	Angkatan 2001-2006	Swasta
14	Teknik	Lulusan 2011 dan 2012	IPK antara 3 sampai 3,5	Angkatan 2001-2006	Swasta
15	Teknik	Lulusan 2011 dan 2012	IPK antara 3 sampai 3,5	Angkatan 2001-2006	Swasta
16	Teknik	Lulusan 2011 dan 2012	IPK antara 3 sampai 3,5	Angkatan 2007-2013	Pemerintah
17	Teknik	Lulusan 2011 dan 2012	IPK lebih dari 3,5	Angkatan 2007-2013	Swasta
18	Teknik	Lulusan 2011 dan 2012	IPK lebih dari 3,5	Angkatan 2007-2013	Swasta
19	Ekonomi	Lulusan 2011 dan 2012	IPK lebih dari 3,5	Angkatan 2007-2013	Swasta
20	Ekonomi	Lulusan 2011 dan 2012	IPK antara 3 sampai 3,5	Angkatan 2007-2013	Swasta
21	Ekonomi	Lulusan 2011 dan 2012	IPK lebih dari 3,5	Angkatan 2007-2013	Swasta
22	Ekonomi	Lulusan 2011 dan 2012	IPK lebih dari 3,5	Angkatan 2007-2013	Swasta
23	Ekonomi	Lulusan 2011 dan 2012	IPK lebih dari 3,5	Angkatan 2007-2013	Pemerintah

TABEL 2. Hasil uji

Node	Atribut	Pemerintah	Swasta	Total	Entropy	Gain	Split Info	Gain Ratio
	Total	72	187	259	0.85270			
1	FAKULTAS			259		0.06549	1.55709	0.04206
	Ekonomi	18	81	99	0.68403			
	Teknik	10	53	63	0.63126			
	Kedokteran	44	53	97	0.99378			
2	TAHUN LULUS			259		0.01625	1.09948	0.01478
	2011-2012	42	84	126	0.91829		0.50571	
	2013-2014	30	99	129	0.78244		0.50084	
	2015-2016	0	4	4	0		0.09292	
3	IPK			259		0.00201	1.39906	0.00143
	<3	7	25	32	0.75784		0.37273	
	3-3.5	37	95	132	0.85586		0.49559	
	>3.5	28	67	95	0.87476		0.53073	
4	ANGKATAN			259		0.00208	0.33445	0.00624
	2001-2006	3	13	16	0.69621		0.2481	
	2007-2013	69	174	243	0.86077		0.08631	

Selanjutnya menghitung nilai *SplitInfo* dari masing-masing nilai atribut dengan menggunakan rumus (3).

$$\text{Gain (Total, Fakultas)} = 0.085270 - \left(\left(\frac{99}{259} \times 0.68403 \right) + \left(\frac{63}{259} \times 0.63126 \right) + \left(\frac{97}{259} \times 0.99378 \right) \right) = 0.06549$$

$$\text{Gain (Total, Tahun Lulus)} = 0.085270 - \left(\left(\frac{126}{259} \times 0.91829 \right) + \left(\frac{129}{259} \times 0.78244 \right) + \left(\frac{4}{259} \times 0 \right) \right) = 0.01625$$

$$\text{Gain Ratio (Fakultas)} = \frac{0.06549}{1.55709} = 0.04206$$

$$\text{Gain (Total, Angkatan)} = 0.085270 - \left(\left(\frac{16}{259} \times 0.68403 \right) + \left(\frac{243}{259} \times 0.86077 \right) \right) = 0.00208$$

$$\text{SplitInfo (IPK, >3)} = \left(\frac{-95}{259} \times \log_2 \left(\frac{95}{259} \right) \right) = 0.53073$$

dan Total dari *SplitInfo* IPK adalah 1.39906

$$\text{SplitInfo (Angkatan, 2001-2006)} = \left(\frac{-16}{259} \times \log_2 \left(\frac{16}{259} \right) \right) = 0.24814$$

$$\text{SplitInfo (Angkatan, 2007-2013)} = \left(\frac{-243}{259} \times \log_2 \left(\frac{243}{259} \right) \right) = 0.08631$$

dan Total dari *SplitInfo* Angkatan adalah 0.33445. Selanjutnya menghitung semua *SplitInfo* dari nilai atribut yang ada seperti cara di atas, sampai mendapatkan nilai Total *SplitInfo* dari nilai atribut yang ada.

$$\text{Gain (Total, IPK)} = 0.085270 - \left(\left(\frac{32}{259} \times 0.75787 \right) + \left(\frac{132}{259} \times 0.85586 \right) + \left(\frac{95}{259} \times 0.87476 \right) \right) = 0.00201$$

$$\text{Gain Ratio (Tahun Lulus)} = \frac{0.01625}{1.09948} = 0.01478$$

$$\text{Gain Ratio (IPK)} = \frac{0.00201}{1.39906} = 0.00143$$

$$\text{Gain Ratio (Angkatan)} = \frac{0.00208}{0.33445} = 0.00624$$

Lakukan perhitungan sampai mendapatkan pohon keputusan. Adapun aturan atau rule yang terbentuk dari pohon keputusan (*decision tree*) terakhir berdasarkan gambar 2 adalah sebagai berikut:

1. *If* Fakultas = Ekonomi *And* Tahun Lulus = Lulusan 2011 dan 2012 *Then* Jenis Pekerjaan = Swasta
2. *If* Fakultas = Ekonomi *And* Tahun Lulus = Lulusan 2013 dan 2014, IPK = IPK antara 3 sampai 3,5, Angkatan = Angkatan 2001-2006 *Then* Jenis Pekerjaan = Pemerintah
3. *If* Fakultas = Ekonomi *And* Tahun Lulus = Lulusan 2013 dan 2014, IPK = IPK antara 3 sampai 3,5, Angkatan = Angkatan 2007-2013 *Then* Jenis Pekerjaan = Swasta
4. *If* Fakultas = Ekonomi *And* Tahun Lulus = Lulusan 2013 dan 2014, IPK = IPK kurang dari 3, Angkatan = Angkatan 2001-2006 *Then* Jenis Pekerjaan = Swasta
5. *If* Fakultas = Ekonomi *And* Tahun Lulus = Lulusan 2013 dan 2014, IPK = IPK kurang

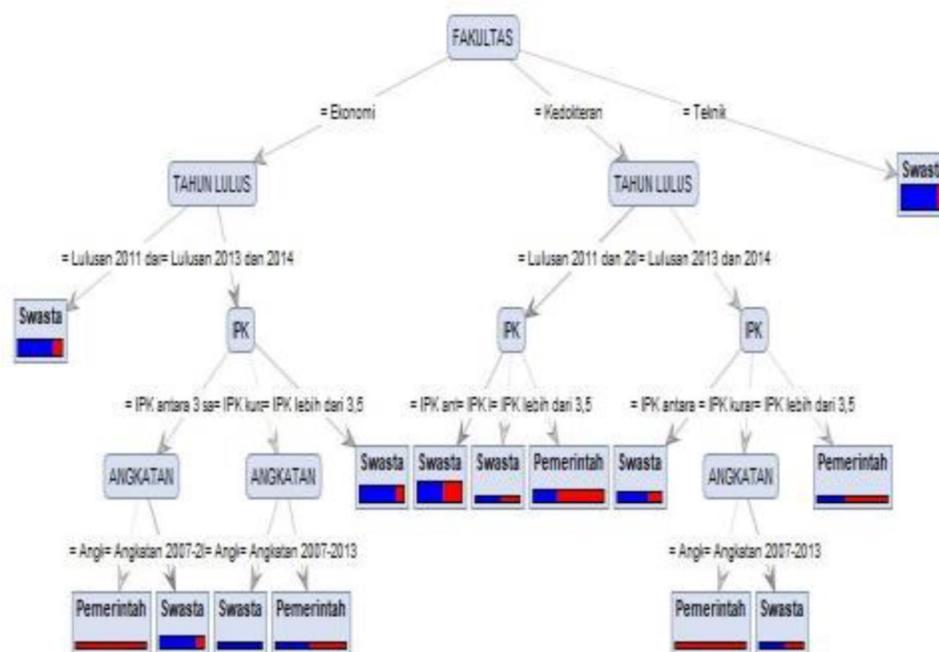
- dari 3, Angkatan = Angkatan 2007-2013
Then Jenis Pekerjaan = Pemerintah
- 6. If Fakultas = Kedokteran And Tahun Lulus = Lulusan 2011-2012, IPK = IPK antara 3 sampai 3.5 Then Jenis Pekerjaan = Swasta
- 7. If Fakultas = Ekonomi And Tahun Lulus = Lulusan 2013 dan 2014, IPK = IPK lebih dari 3,5 Then Jenis Pekerjaan = Swasta
- 8. If Fakultas = Kedokteran And Tahun Lulus = Lulusan 2011-2012, IPK = IPK kurang dari 3 Then Jenis Pekerjaan = Swasta
- 9. If Fakultas = Kedokteran And Tahun Lulus = Lulusan 2011-2012, IPK = IPK lebih dari 3,5 Then Jenis Pekerjaan = Swasta
- 10. If Fakultas = Kedokteran And Tahun Lulus = Lulusan 2013-2014, IPK = IPK antara 3 sampai 3,5 Then Jenis Pekerjaan = Swasta
- 11. If Fakultas = Kedokteran And Tahun Lulus = Lulusan 2013-2014, IPK = IPK kurang dari 3, Angkatan = Angkatan 2001-2006 Then Jenis Pekerjaan = Pemerintah
- 12. If Fakultas = Kedokteran And Tahun Lulus = Lulusan 2013-2014, IPK = IPK kurang dari 3, Angkatan = Angkatan 2007-2013 Then Jenis Pekerjaan = Swasta
- 13. If Fakultas = Kedokteran And Tahun Lulus = Lulusan 2013-2014, IPK = IPK lebih dari 3,5 Then Jenis Pekerjaan = Pemerintah
- 14. If Fakultas = Teknik Then Jenis Pekerjaan = Swasta

Pengujian Data Menggunakan Software RapidMiner. Pada Gambar 2 adalah hasil proses

pengujian yang dilakukan menggunakan software RapidMiner.

Berdasarkan pohon keputusan pada Gambar 2 dapat dilihat bahwa pengaruh yang dominan terhadap jenis pekerjaan alumni adalah Fakultas karena Fakultas dijadikan sebagai akar atau *root* dari pohon keputusan tersebut. Yang dapat menjadi sebuah akar pada pohon keputusan adalah atribut yang memiliki nilai *gain ratio* tertinggi dibandingkan dengan atribut yang lainnya.

Sedangkan pertimbangan terakhir yang berpengaruh pada jenis pekerjaan adalah Angkatan karena berada pada *node* terakhir. Gambar 2. menunjukkan deskripsi lengkap tentang pohon keputusan (*decision tree*) yang terbentuk dari 259 data alumni yang berasal dari 3 fakultas yaitu fakultas Ekonomi, fakultas Teknik, dan fakultas Kedokteran untuk menentukan jenis pekerjaan pemerintah dan swasta. Berdasarkan hasil pengujian yang telah dilakukan pada data alumni telah diperoleh sebuah klasifikasi jenis pekerjaan alumni. Dan hasil dari pengujian tersebut dapat digunakan sebagai informasi strategis yang dapat dijadikan sebagai sebuah pengetahuan. Sehingga pengetahuan ini dapat dijadikan sebagai pendukung keputusan atau suatu kebijakan strategis untuk Universitas Muhammadiyah Yogyakarta



GAMBAR 2. Pohon Keputusan

KESIMPULAN

Berdasarkan uraian pada bab-bab sebelumnya yang telah dibahas, maka dapat ditarik kesimpulan sebagai berikut:

1. Berdasarkan pohon keputusan (*decision tree*) atribut yang paling berpengaruh terhadap jenis pekerjaan alumni adalah Fakultas. Hal itu terjadi karena fakultas sebagai akar dari pohon keputusannya, dan memiliki nilai *gain ratio* yang paling tinggi di antara atribut yang lain, selain atribut fakultas masih ada atribut lain yang berpengaruh yaitu tahun lulus, dan ipk. Sedangkan angkatan tidak terlalu berpengaruh terhadap jenis pekerjaan alumni.
2. Berdasarkan hasil penelitian penerapan metode algoritma C4.5 cocok diterapkan pada proses klasifikasi jenis pekerjaan alumni fakultas Ekonomi, fakultas Teknik, dan fakultas Kedokteran.
3. Pada *dataset* jenis pekerjaan alumni telah diterapkan salah satu metode *data mining* yaitu klasifikasi menggunakan algoritma C4.5 dan telah dilakukan pengujian menggunakan *Software RapidMiner* menghasilkan klasifikasi yang sama.

UCAPAN TERIMA KASIH

Penulis dapat menuliskan ucapan terima kasih kepada CDC UMY dan LP3M UMY yang telah membantu dalam penelitian yang dipublikasikan dalam jurnal ini.

DAFTAR PUSTAKA

- Andriani, Anik. (2012). Penerapan Algoritma C4. 5 Pada Program Klasifikasi Mahasiswa Dropout. In , 139–47.
- Elisa, Erlin. (2017). Analisa dan Penerapan Algoritma C4.5 Dalam Data Mining Untuk Mengidentifikasi Faktor-Faktor Penyebab Kecelakaan Kerja Kontruksi PT.Arupadhatu Adisesanti. Jurnal Online Informatika 2 (1): 36–41.
- Fansuri, Farid. (2012). Jurnal Komputer dan Informatika (KOMPUTA). 1: 5.

Ginting, Selvia Lorena Br, Wendi Zarman, and Ida Hamidah. (2014). Analisis Dan Penerapan Algoritma C4. 5 Dalam Data Mining Untuk Memprediksi Masa Studi Mahasiswa Berdasarkan Data Nilai Akademik. No. November.

Kamagi, David Hartanto, and Seng Hansun. (2014). Implementasi Data Mining Dengan Algoritma C4. 5 Untuk Memprediksi Tingkat Kelulusan Mahasiswa. Vol. VI, no. 1: 15–20.

Mabrur, Angga Ginanjar. (2012). Penerapan Data Mining Untuk Memprediksi Kriteria Nasabah Kredit. Komputa: Jurnal Ilmiah Komputer Dan Informatika 1 (1).

Nugroho, Yusuf Sulistyono. (2014). Penerapan Algoritma C4. 5 Untuk Klasifikasi Predikat Kelulusan Mahasiswa Fakultas Komunikasi Dan Informatika Universitas Muhammadiyah Surakarta. In , A1–6.

Santoso, Teguh Budi. (n.d). Analisa Dan Penerapan Metode C4.5 Untuk Prediksi Loyalitas Pelanggan. 10: 10.

Sembiring, Muhammad Ardiansyah, Mustika Fitri Larasati Sibuea, and Andy Saptana. (n.d). Analisa Kinerja Algoritma C.45 Dalam Memprediksi Hasil Belajar, 7.

Widayu, Hikma, Surya Darma Nasution, Natalia Silalahi, and Mesran Mesran. (2017). Data Mining Untuk Memprediksi Jenis Transaksi Nasabah Pada Koperasi Simpan Pinjam Dengan Algoritma C4. 5. Media Informatika Budidarma 1 (2).

PENULIS:

Asroni

Program Studi Teknik Informatika, Fakultas Teknik, Universitas Muhammadiyah Yogyakarta, D.I. Yogyakarta

Email: asroni@umy.ac.id

Badrahini Masajeng Respati

Program Studi Teknik Informatika, Fakultas Teknik, Universitas Muhammadiyah Yogyakarta, D.I Yogyakarta

Email: masajeng@umy.ac.id

Slamet Riyadi

Profram Studi Teknik Informatika, Fakultas
Teknik, Universitas Muhammadiyah
Yogyakarta, Bantul, D.I Yogyakarta

Email: riyadi@umy.ac.id