

KLASIFIKASI KOMPTENSI MAHASISWA DENGAN ALGORITMA DECISION TREE DALAM MENETUKAN KELAYAKAN MATA KULIAH KOSENTRASI

Suhardi Rustam

Fakultas Ilmu Komputer , Universitas Ichsan Gorontalo

Jl. Achmad Nadjamuddin No. 17 Telp. 0435-829975 Fax. 0435-829976 Kota Gorontalo

suhardirstm@gmail.com

Permasalahan yang selalu menjadi rutinitas yang di alami jurusan dalam fakultas adalah mengelompokan mahasiswa yang akan memilih mata kuliah kosentrasi, mata kuliah kosentrasi adalah kelompok mata kuliah jurusan yang akan memfokuskan mahasiswa dalam menekuni satu topik keilmuan dan akan berlanjut ke topik penelitian dalam penyelesaian tugas akhir. Komptensi yang diasah dalam mata kuliah kosentrasi adalah sesuai dengan kemampuan yang diperoleh dari mata kuliah yang telah diperoleh dari mata kuliah di semester sebelumnya, kekurangtepatan dalam mengkalisifikasi mahasiswa ikut mempengaruhi proses optimalisasi komptensi dan lulus tidak tepat waktu, Algoritma Decision tree yang memiliki karakteristik percabangan akar (tree). eksperimen ini menggunakan tools rapidminer dalam memproses model klasifikasi berbasis decision tree dengan algoritma C4.5 dari dataset yang telah dipreprocessing yang berfungsi untuk membersihkan missing value dan outlier, hasil pengklasifikasian kompetensi dengan algoritma decision tree menggunakan algoritma C4.5 adalah dengan hasil akurasi 89.86%, classification_error 10.14% dan weight mean recall 90.34%. hasil akurasi yang didapatkan merupakan hasil yang terbaik

Kata Kunci : Data Mining Akademik, Kompetensi, Decision Tree, Algoritma C4.5

I. PENDAHULUAN

Komptensi merupakan kemampuan maksimal yang akan diperoleh mahasiswa selama proses perkuliahan akademik, pengelolaan kemampuan dalam manajemen akademik menjadi permasalahan yang menjadi titik fokus fakultas pada universitas dalam mencapai visi misi organisasi yang akan mencapai lulus yang memiliki komptensi yang maksimal, kekurang maksimalnya dalam pengelompokan mahasiswa yang kelompok mata kuliah penjurusan akan mempengaruhi optimasi proses pembelajaran akademik, misalnya minat, motivasi, kurang fokus dan kurangnya perhatian dalam melakukan

eksperimen terhadap tugas.

Kurang tepatnya mahasiswa dalam menentukan mata kuliah kosentrasi juga akan mempengaruhi target lulusan yang tidak tepat waktu sehingga membentuk kelompok baru sekaligus masalah yang akan menumpuk pada data administiras jurusan dan juga pada data akademik mahasiswa. Rekaman data akademik sebelum mata kuliah kosentrasi perlu dioptimalkan sehingga memunculkan pola pengetahuan yang memungkinkan menjadi alternatif dalam mengatasi permasalahan pemilihan komptensi.

Pengklasifikasian komptensi menjadi kebutuhan utama dalam meningkat performa jurusan/fakultas dalam mengambil keputusan berbasis kompetensi sesuai dengan visi misi. Nilai akademik mahasiswa menjadi informasi awal dalam mendalami komptensi mahasiswa.

Rekaman data akademik yang akan digunakan yaitu data mahasiswa angkatan 2016-2017 yang telah mengambil kelas mata kuliah kosentrasi. Dan data ini juga sebagai data testing yang akan dicoba kedalam model algoritma klasifikasi. Rekaman data akademik ini akan digunakan untuk memberikan pola permasalahan dengan menggunakan algoritma data mining. Algoritma yang digunakan dalam permasalahan ini adalah algoritma decision tree untuk membentuk pola klasifikasi komptensi mahasiswa pada kelompok mata kuliah kosentrasi jurusan.

Adanya kemudahan dalam melakukan penilaian terhadap rekaman nilai akademik harus memberikan informasi dan mendukung keputusan (Decission Support System) penilaian terhadap rekaman data akademik inputan nilai akademik yang nantinya akan dijadikan acuan kompetensi[1] dari mahasiswa.

Klasifikasi merupakan salah satu teknik dalam data mining. Klasifikasi[2] didefinisikan sebagai berikut: (1) Meramalkan kategori label kelas (diskrit atau nominal). (2) Mengklasifikasikan (membuat suatu model) berdasarkan himpunan pelatihan dan nilai-nilai (label kelas) dalam suatu atribut klasifikasi dan menggunakannya di dalam mengklasifikasikan data baru. Klasifikasi[3] adalah proses untuk menemukan model atau

fungsi yang menjelaskan atau membedakan konsep atau kelas data, dengan tujuan untuk dapat memperkirakan kelas dari suatu objek yang labelnya tidak diketahui. Model itu sendiri dapat berbentuk pohon pengambilan keputusan (decision tree)

Decision Tree (pohon keputusan)[4] adalah flowchart seperti struktur tree, di mana tiap internal node menunjukkan sebuah test pada sebuah atribut, tiap cabang menunjukkan hasil dari test, dan leaf node menunjukkan class-class atau class. Decision Tree berusaha menemukan atribut yang tepat untuk menyelesaikan dan menentukan kelas. Dengan kata lain, atribut yang berada di posisi paling atas dalam decision tree[5] adalah atribut yang paling berpengaruh dalam menentukan hasil prediksi. Oleh karena itu, proses pembuatan model decision tree merupakan proses interative untuk menilai atribut paling berpengaruh yang akan menjadi percabangan dari struktur pohon. Algoritma C4.5[6] merupakan algoritma klasifikasi dengan teknik Decision tree (pohon keputusan) yang terkenal dan disukai karena memiliki kelebihan-kelebihan. Kelebihannya adalah dapat mengolah data numerik (kontinyu) dan diskret, dapat menangani nilai atribut yang hilang, dapat menghasilkan aturan aturan yang mudah diinterpretasikan dan tercepat diantara algoritma- algoritma yang lain.

Dalam penelitian ini untuk proses pada decision tree adalah mengubah bentuk data (tabel) menjadi model pohon, mengubah model pohon menjadi rule, dan menyederhanakan rule. Sehingga hasil terbentuk adalah pola klasifikasi kompetensi mahasiswa sebagai rekomendasi keputusan untuk kelompok mata kuliah konsentrasi jurusan

II. LANDASAN TEORI

A. Data Mining

data mining atau penambangan data dapat didefinisikan sebagai proses seleksi, eksplorasi, dan pemodelan dari sejumlah besar data untuk menemukan pola atau kecenderungan yang biasanya tidak disadari keberadaannya. Data mining dapat dikatakan sebagai proses mengekstrak pengetahuan dari sejumlah besar data yang tersedia. Pengetahuan yang dihasilkan dari proses data mining harus baru, mudah dimengerti, dan bermanfaat. Dalam data mining, data disimpan secara elektronik dan diproses secara otomatis oleh komputer menggunakan teknik dan perhitungan tertentu[7].

B. Algoritma Decision Tree

Algoritma decision tree merupakan salah satu algoritma yang digunakan untuk melakukan klasifikasi atau pengelompokan dan bersifat prediktif. Dasar algoritma decision tree adalah pembentukan pohon keputusan. Proses pada pohon keputusan adalah mengubah bentuk data (tabel) menjadi model pohon, mengubah model pohon menjadi rule dan menyederhanakan rule [8].

Tahapan Algoritma Decision Tree Secara umum algoritma C4.5 untuk membangun pohon keputusan adalah sebagai berikut:

1. Siapkan *data training*,
2. Hitung nilai *entropy*, menghitung dengan rumus sebagai berikut

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

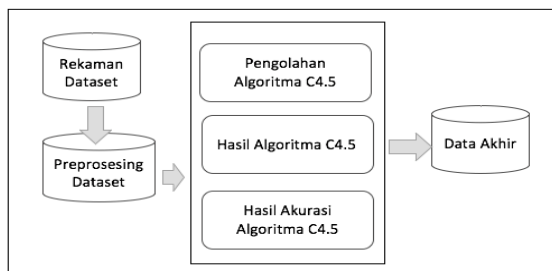
3. Lanjut hitung nilai *gain* dalam setiap atribut, nilai tertinggi sebagai nilai akar,

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

III. METODE PENELITIAN

Penggunaan metode penelitian ini, mempergunakan cara penelitian eksperimen, untuk tahap penelitian dimulai dari observasi, pengumpulan rekaman dataset akademik mahasiswa sampai dengan mendapatkan hasil evaluasi dan hasil untuk mendapatkan tujuan yang akan dicapai, adapun tahapannya adalah, di mulai dari penyiapan rekaman dataset kemudian dilakukan langkah preprocessing terhadap data[9] untuk menghilangkan outliernya, selanjutnya adalah menormalisasikan data kemudian menggunakan *tool data mining rapidminer* untuk mengelola data tersebut dengan menggunakan metode Algoritma Decision Tree yaitu Algoritma C4.5. Model yang penelitian ini adalah sebagai berikut :

A. Eksperimen



Gambar 1. Model Ekperimen Algoritma C4.5

B. Observasi dan Pengumpulan Data

Dalam penelitian ini, rekaman dataset diperoleh dari penelitian yang dilakukan pada data universitas ichsan Gorontalo yang berkaitan langsung dengan bagian pusat pengelolaan data kemahasiswaan universitas (BAK Universitas) dengan fokus rekaman dataset fakultas ilmu komputer. Identifikasi variable[10] dalam penelitian menentukan cara pengolahan data dan alat yang digunakan untuk menganalisis data, adapun variable dalam penelitian ini adalah sebagai berikut :

- a. Nilai mata kuliah Jurusan, adalah nilai mata kuliah yang utam untuk mata kuliah wajib jurusan
- b. Nilai mata kuliah prasyarat konsentrasi, adalah nilai mata kuliah yang pelengkap untuk mata kuliah konsentrasi

c. Jumlah kehadiran dari mata kuliah prasyarat konsentrasi, merupakan nilai kehadiran mahasiswa dalam mengikuti mata kuliah tersebut

C. Pengolahan Data Awal

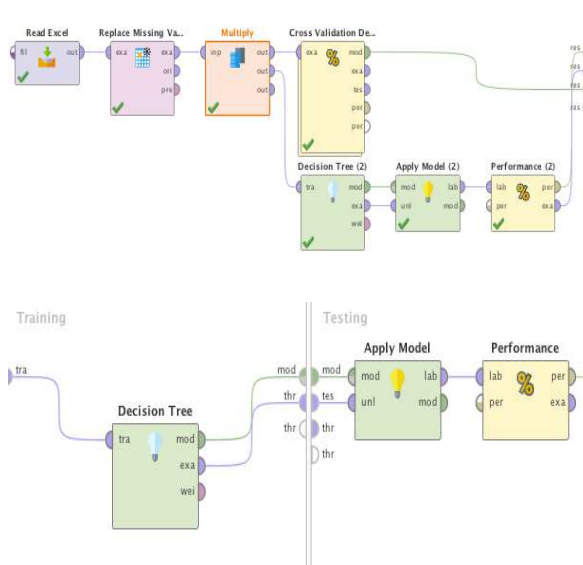
Data awal penelitian ini perlu dibatasi pada tahapan ini melakukan *preprocessing* terhadap data untuk menghilangkan outlier, membersihkan setiap variable yang terkait sehingga untuk ke tahap penggunaan selanjutnya telah siap

D. Eksperimen

Ekperimen pada tahapan ini adalah melakukan proses data mining[11], menentukan data training, data testing, proses data training dan proses data testing. Pada proses data mining yaitu pengolahan dataset untuk menemukan pola informasi dan pengetahuan dengan algoritma C4.5 dengan menggunakan tools rapidminer.

E. Evaluasi

Dalam pengujian eksperimen yang dilakukan berupa eksperimen dataset dalam mengklasifikasi sehingga membentuk pohon klasifikasi dengan nilai akurasi sebagai nilai performance pemodelan algoritma decision tree berdasarkan nilai kriteria.



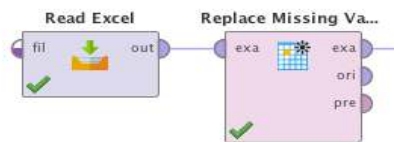
Gambar 3, Pemodelan Decision Tree C4.5

C. Hasil Testing

IV. HASIL DAN PEMBAHASAN

A. Model Processing

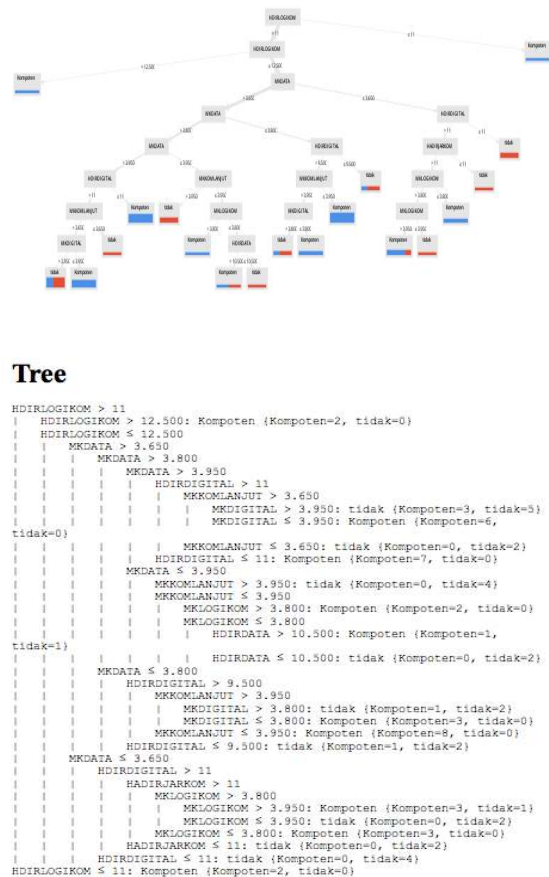
Dalam menentukan data training dan data testing dengan jumlah sampel populasi (n) sebanyak 70 record populasi dan sebagai sampel sebanyak 30 sampel. Pada model processing adalah model yang dibentuk untuk membersihkan missing value dan outlier untuk mengolah data selanjutnya, adapun modelnya adalah sebagai berikut :



Gambar 2 , Model Processing

B. Proses training dan Testing

Pada prose training dan Testing, dilakukan proses pemodelan eksperimen dengan menggunakan tools rapidminer seperti pada Gambar 3.



Gambar 4, Pohon Klasifikasi Algoritma C4.5

D.Hasil Eksperimen dan Pembahasan

Setelah melakukan eksperimen dataset sejumlah dataset 70 record (sampel), eksperimen ini menggunakan *tools rapidminer* dalam memproses model klasifikasi berbasis *decision tree* dengan algoritma C4.5 dari dataset yang telah dipreprocessing yang berfungsi untuk membersihkan *missing value* dan *outlier* maka didapatkan hasil akurasi dalam tabel dibawah ini,

Table 1, Hasil Eksperimen dengan Algoritma C4.5

No	Item	keterangan
1.	Model Algoritma	Decision tree Algoritma C4.5
2.	Accuracy	Hasil : 89.86%
3.	classification error	Hasil : 10.14%
4.	weighted_mean_recall	Hasil : 90.34%,

Pada hasil eksperimen pada tabel 1, merupakan hasil eksperimen dengan menggunakan model eksperimen klasifikasi *decision tree* dengan algoritma C4.5, dengan hasil akurasi 89.86%, *classification_error* 10.14% dan *weight mean recall* 90.34%

V. KESIMPULAN DAN SARAN

Berdasarkan hasil eksperimen pada bagian pembahasan sebelumnya dengan model tersebut, pengklasifikasian kompetensi dengan algoritma *decision tree* menggunakan algoritma C4.5 adalah dengan hasil akurasi 89.86%, *classification_error* 10.14% dan *weight mean recall* 90.34%. hasil akurasi yang didapatkan merupakan hasil yang terbaik. Dan adapun saran dalam penelitian ini, untuk keperluan lebih lanjut untuk algoritma C4.5 data mining khususnya bagian klasifikasi dalam pengolahan data akademik mahasiswa berbasis *decision tree*, maka disarankan untuk menguji coba perbandingan algoritma klasifikasi yang lain dikarenakan dalam eksperimen ini fokus menggunakan algoritma C4.5, penelitian semacam ini dapat dikembangkan untuk mendapatkan hasil yang optimal di masing-masing algoritma klasifikasi.

UCAPAN TERIMA KASIH

Tulisan ini, saya sebagai ucapan terima kasih kepada Lembaga Lemlit sebagai wadah yang memotivasi dan mendukung untuk meneliti dan mempublikasi artikel, rekan-rekan dosen di Fakultas Ilmu Komputer Universitas Ichsan Gorontalo atas saran dan dukungannya

DAFTAR PUSTAKA

[1] A. S. Sunge, "Prediksi Kompetensi Karyawan

Menggunakan Algoritma C4 . 5 (Studi Kasus : PT Hankook Tire Indonesia)," *Semin. Nas. Teknol. Inf. dan Komun. 2018 (SENTIKA 2018)*, vol. 2018, no. Sentika, pp. 23–24, 2018.

[2] R. T. Jurnal, "Pembentukan Model Klasifikasi Data Lama Studi Mahasiswa Stmik Indonesia Menggunakan Decision Tree Dengan Algoritma Nbtrees," *Kilat*, vol. 6, no. 2, pp. 145–148, 2018, doi: 10.33322/kilat.v6i2.135.

[3] N. Nafi'iyah, "Algoritma Cart Dalam Penentuan Pohon Keputusan," vol. 7, no. 2, 2015.

[4] M. H. Darmawan and R. E. Indrajit, "Implementasi Konsep Business Intelligence Untuk Menentukan Kebutuhan Training Pada Klien," no. November, pp. 1–2, 2017.

[5] H. Susanto and S. Sudiyatno, "Data mining untuk memprediksi prestasi siswa berdasarkan sosial ekonomi, motivasi, kedisiplinan dan prestasi masa lalu," *J. Pendidik. Vokasi*, vol. 4, no. 2, pp. 222–231, 2014, doi: 10.21831/jpv.v4i2.2547.

[6] U. Pujiyanto and P. Y. Ristanti, "Perbandingan kinerja metode C4.5 dan Naive Bayes dalam klasifikasi artikel jurnal PGSD berdasarkan mata pelajaran," *Tekno*, vol. 29, no. 1, p. 50, 2019, doi: 10.17977/um034v29i1p50-67.

[7] N. Walidaini, M. A. Mukid, A. Prahutama, and A. Rusgiyono, "Analisis Diskriminan Berganda Dengan Peubah Bebas Campuran Kategorik Dan Kontinu Pada Klasifikasi Indeks Prestasi Kumulatif Mahasiswa," *Media Stat.*, vol. 10, no. 2, p. 71, 2017, doi: 10.14710/medstat.10.2.71-83.

[8] A. R. Khadafy, "Penerapan Naive Bayes untuk Mengurangi Data Noise pada Klasifikasi Multi Kelas dengan Decision Tree," *J. Intell. Syst.*, vol. 1, no. 2, pp. 136–142, 2015.

[9] S. Defit and G. W. Nurcahyo, "Penerapan Algoritma C4 . 5 Pada NUPTK Untuk Menentukan Pola Sertifikasi Guru Dengan Menggunakan Metode Klasifikasi Decision Tree (Studi Kasus : Pada Dinas Pendidikan Pemuda dan Olahraga Kabupaten Rokan Hulu) Of Computer Science," *J. Comput. Sci.*, vol. 1, no. 1, pp. 69–83, 2015.

[10] F. I. Komputer and U. D. Nuswantoro, "Menggunakan Metode Decision Tree Pada," vol. 8, no. 1, pp. 1–8, 2010.

[11] A. E. Pramadhani and T. Setiadi, "Penerapan Data Mining untuk Klasifikasi Penyakit ISPA dengan Algoritma Desicion Tree," *J. Sarj. Tek. Inform. e-ISSN 2338-5197*, vol. 2, no. 1, pp. 831–839, 2014.