

PERANCANGAN SISTEM MENGGUNAKAN ALGORITMA C 4.5 UNTUK MEMPREDIKSI JUMLAH MAHASISWA MENGAMBIL MATAKULIAH SYARAT PLK

Sonia Afriliani¹⁾, Khairudin²⁾, Karmila Suryani²⁾

¹⁾Mahasiswa Program Studi PTIK FKIP Universitas Bung Hatta Padang

²⁾Dosen Program Studi PTIK FKIP Universitas Bung Hatta Padang

E-mail : sonia.afriliani10@gmail.com

ABSTRAK

Each semester the department only estimate how many classes will be opened in accordance with how many students manually, in which there is no system to determine the number of students taking main subject for education field practice condition, which led to the error of a data include class density and shortage of capacity class, even classes that previously united with the same class that causes students confused by the changes that occur, to avoid problems occur then the researchers conducted a study using data mining to predict the number of students who taking an education field practice with the assistance of C4.5 algorithm as a procedure makes the tree determinant of the decision, which is to find students who have or have not taken the previously education field practice requirements and how classes will be opened to the course requirements of education field practice which are Pengantar Pendidikan, Profesi Pendidikan, Algoritma and Pemrograman, Paket Aplikasi and Media Pembelajaran Berbasis TIK

Keywords: *Data Mining, Subjects for Education Field Practice Condition, Algorithm C4.5*

PENDAHULUAN

Perkembangan teknologi saat ini sangatlah pesat, banyak *programmer* yang merancang sistem informasi dengan berbagai aplikasi dan teknik pembuatannya. Akan tetapi masih belum banyak *programmer* menciptakan sistem informasi menggunakan algoritma C4.5 begitu juga di lingkungan Universitas Bung Hatta, salah satunya program studi Pendidikan Teknik Informatika dan Komputer.

Pendidikan Teknik Informatika dan Komputer (PTIK) berdiri pada tanggal 24 maret 2011 yang mana merupakan salah

satu jurusan yang ada di Universitas Bung Hatta. PTIK memiliki beban sks sebanyak 150 yang harus ditempuh oleh mahasiswa diantaranya yaitu Matakuliah Pengembangan Kepribadian (MPK), Matakuliah Keilmuan dan Keterampilan (MKK), Matakuliah Perilaku Berkarya (MPB), dan Matakuliah Berkehidupan Bersama (MBB).

Adapun matakuliah yang wajib ditempuh oleh mahasiswa sebagai syarat Praktek Lapangan Kerja (PLK) baik semester ganjil dan semester genap dari MPK, MKK, MPB, dan MBB. Setiap semester pihak jurusan memperkirakan

berapa kelas yang akan dibuka. Serta cara melakukan pengumpulan data mahasiswa untuk mengambil matakuliah dan yang belum mengambil matakuliah belum terkomputerisasi, sehingga dapat terjadi kekeliruan terhadap pihak jurusan untuk menentukan berapa lokal yang dibuka tiap semesternya.

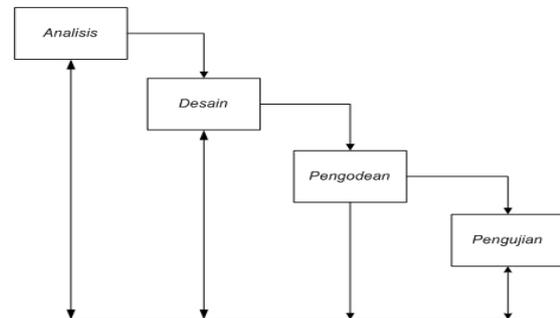
Pada saat mengambil matakuliah bisa terjadi kepadatan kelas yang membuat mahasiswa ada yang terpaksa mengambil matakuliah yang seharusnya belum dipilih karena kelas yang diambil kapasitasnya sudah penuh, juga bisa terjadi kekurangan kapasitas kelas disebabkan kelas yang dibuka tidak sesuai dengan kapasitas mahasiswa, dan kelas yang kekurangan kapasitas disatukan dengan kelas kapasitasnya belum penuh. Sehingga terjadi perubahan hari, jadwal, kelas dan dosen yang ngajar, oleh sebab itu mahasiswa merasa bingung dengan perubahan kelas yang membuat ketidaknyamanan bagi mahasiswa disebabkan belum ada sistem untuk menentukan berapa banyak mahasiswa mengambil matakuliah.

METODE PENELITIAN

Analisis dan Perancangan Sistem

Sistem ini menggunakan metode *waterfall* dalam perancangan sistem. Inti dari metode *waterfall* adalah pengerjaan

dari suatu sistem dilakukan secara berurutan atau secara linear. Berikut adalah gambar model *waterfall*



Gambar 1 Metode Pengembangan Sistem Waterfall (Rosa A. S, 2013:28)

a. Analisis Data

1. Semester variabel semester berisi data semester mahasiswa yang telah ditempuh, yang mana semester ganjil yaitu semester 1, 3, dan 5 yang akan di tempuh mahasiswa.
2. Matakuliah variabel ini berisi matakuliah yang telah diambil sebelumnya atau matakuliah yang belum diambil oleh mahasiswa.
3. Jenis kelamin variabel ini berisi data jenis kelamin dari setiap mahasiswa yang akan digunakan sebagai pembentuk keputusan. Pengelompokan yang ada berdasarkan ketentuan yang dibuat program adalah laki-laki dan perempuan.
4. Nilai variabel berisi seluruh kemungkinan nilai setiap matakuliah yang dimiliki oleh mahasiswa untuk diisi pada proses input program. Nilai

yang sudah ditentukan pada program ini antara lain A, A-, B+, B, B-, C+, C, D, dan E.

5. Keputusan variabel keputusan merupakan data yang berfungsi untuk menentukan hasil keputusan. Dalam pengelompokan data sudah ditentukan secara tetap agar tidak terjadi kesalahan dalam perhitungan proses program. Data keputusan hanya memiliki dua buah nilai yaitu “Ya” dan “Tidak”.

Tabel 1 Atribut Masukan

Atribut	Keterangan
Semester	Ganjil (1, 3 dan 5)
Matakuliah semester 1 syarat PL	Pengantar Pendidikan
Matakuliah semester 3 syarat PL	Profesi Pendidikan, Algoritma dan Pemrograman 2
Matakuliah semester 5 syarat PL	Paket Aplikasi, Media Pembelajaran Berbasis TIK
Jenis Kelamin	Laki-laki, Perempuan
Nilai	A, A-, B+, B, B-, C+, C, D, E

Tabel 3.1 Perhitungan Node 1

Baris total kolom *entropy* pada tabel 3.2 dihitung dengan rumus Persamaan 2, sebagai berikut:

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

Entropy(total)

$$= \left(-\frac{122}{150} * \log_2 \left(\frac{122}{150} \right) \right) + \left(-\frac{28}{150} * \log_2 \left(\frac{28}{150} \right) \right)$$

Entropy(total) = 0,69444607

Sedangkan nilai Gain pada baris dihitung menggunakan rumus persamaan 1, sebagai berikut:

Gain(total, atribut)

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

Gain(total, atribut)

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

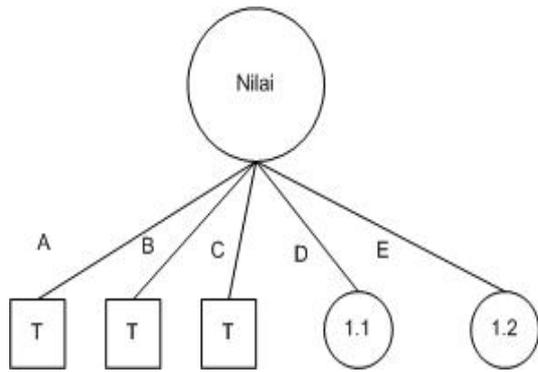
Gain (total, atribut) = 0,69444607 -

$$\left(\left(\frac{30}{150} \times 0 + \frac{60}{150} \times 0,3533 + \frac{60}{150} \times 0,97095 \right) \right)$$

Sehingga didapat *gain (total, atribut) = 0,161154834*, *gain* yang tertinggi jadi akar.

pohon keputusan. Pohon

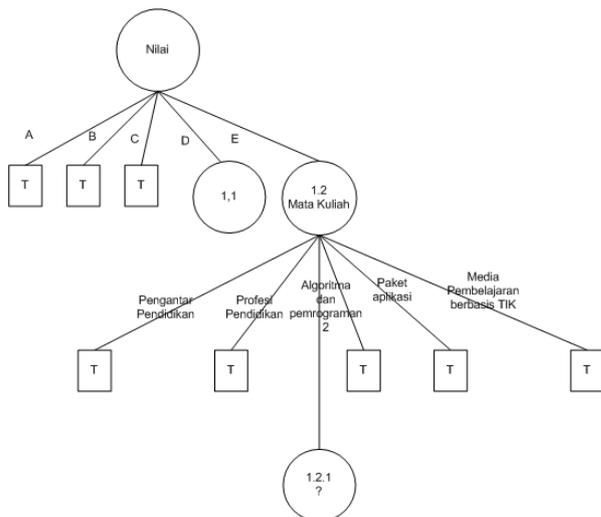
keputusan hasil perhitungan node 1 seperti dibawah ini:



Gambar 2 Pohon Keputusan Hasil Perhitungan Node 1

1. Menghitung jumlah kasus nilai E, jumlah kasus untuk keputusan ya, dan jumlah kasus untuk keputusan tidak dan nilai *entropy* dari setiap kasus yang dibagi berdasarkan atribut, tahun, semester, jenis kelamin, dan matakuliah syarat PLK. Setelah itu melakukan perhitungan *Gain* untuk setiap atribut.

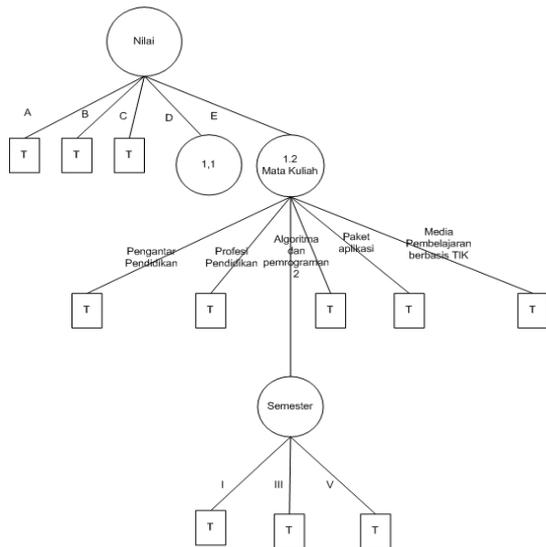
Pohon keputusan hasil perhitungan node 1.2 seperti dibawah ini:



Gambar 3 Pohon Keputusan Hasil Perhitungan Node 1.2

Pada gambar 3 terlihat nilai sebagai akar dan matakuliah sebagai batang yang mana terdiri dari ranting-rantingnya yaitu keputusan ya dan keputusan tidak, keputusan terlihat dari banyaknya keputusan ya dan tidak jika keputusannya lebih banyak ya maka rantingnya ya, dan jika keputusannya lebih banyak tidak maka keputusannya tidak sebagai rantingnya.

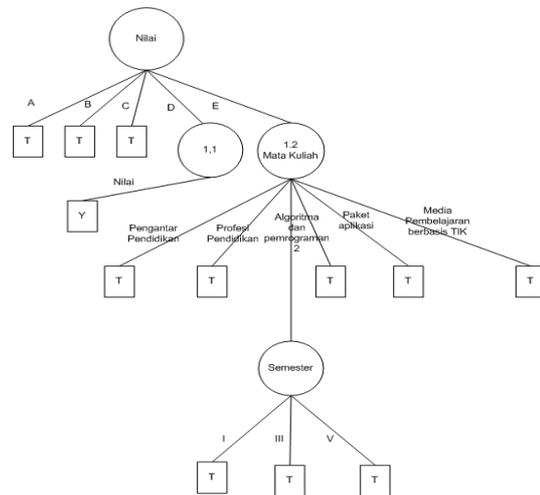
2. Menghitung jumlah kasus matakuliah, jumlah kasus untuk keputusan ya, dan jumlah kasus untuk keputusan tidak dan nilai *entropy* dari setiap kasus yang dibagi berdasarkan atribut, tahun, semester, jenis kelamin, dan semester syarat PLK. Setelah itu melakukan perhitungan *Gain* untuk setiap atribut. hasil gain yang paling tertinggi yaitu Nilai, sehingga dijadikan sebagai akar dari pohon keputusan. Pohon keputusan hasil perhitungan node 1.2.1 seperti dibawah ini:



Gambar 4 Pohon Keputusan Hasil Perhitungan Node 1.2.1

Pada gambar 3 terlihat nilai sebagai akar dan matakuliah sebagai batang yang mana terdiri dari ranting-rantingnya yaitu keputusan ya dan keputusan tidak, keputusan terlihat dari banyaknya keputusan ya dan tidak jika keputusannya lebih banyak ya maka rantingnya ya, dan jika keputusannya lebih banyak tidak maka keputusannya tidak sebagai rantingnya, dan terdapat semester sebagai batang dari node 1.2.1. Pada pohon keputusan untuk 1.1, pada nilai D tidak ditemukan berapa jumlah mahasiswa yang di peroleh oleh mahasiswa sehingga dapat disimpulkan pada gambar dibawah ini, serta

keseluruhan pohon pada kasus yang diteliti oleh peneliti.



Gambar 5 Pohon Keputusan Hasil Perhitungan Node 1.1

Hasil dan PEMBAHASAN

a. Hasil Penelitian



Gambar 6 tampilan lakukan Mining C4.5

Setelah lakukan Mining C4.5, maka akan terlihat pohon keputusan pada saat diklik tombol Pohon Keputusan C4.5 seperti berikut.

```

Pohon Keputusan:
nilai = A (Tidak = 53, Ya = 0) : Tidak
nilai = A- (Tidak = 72, Ya = 0) : Tidak
nilai = B+ (Tidak = 32, Ya = 0) : Tidak
nilai = B (Tidak = 41, Ya = 0) : Tidak
nilai = B- (Tidak = 23, Ya = 0) : Tidak
nilai = C+ (Tidak = 16, Ya = 0) : Tidak
nilai = C (Tidak = 3, Ya = 0) : Tidak
nilai = D (Tidak = 0, Ya = 0) :
nilai = E (Tidak = 7, Ya = 53) : ?
| mk_prasyarat = Pengantar Pendidikan (Tidak = 0, Ya = 0) : Ya
| mk_prasyarat = Profesi Pendidikan (Tidak = 3, Ya = 2) : Ya
| mk_prasyarat = Algoritma dan Pemrograman 2 (Tidak = 3, Ya = 2) : ?
| | tahun = 2012 (Tidak = 0, Ya = 1) : Ya
| | tahun = 2013 (Tidak = 3, Ya = 1) : Tidak
| mk_prasyarat = Paket Aplikasi (Tidak = 0, Ya = 19) : Ya
| mk_prasyarat = Media Pembelajaran Berbasis TIK (Tidak = 1, Ya = 30) : Ya

Rule:
1. if (nilai == A) then Tidak (id = 1)
2. if (nilai == A-) then Tidak (id = 2)
3. if (nilai == B+) then Tidak (id = 3)
4. if (nilai == B) then Tidak (id = 4)
5. if (nilai == B-) then Tidak (id = 5)
6. if (nilai == C+) then Tidak (id = 6)
7. if (nilai == C) then Tidak (id = 7)
8. if (nilai == D) then (id = 8)
9. if (nilai == E AND mk_prasyarat == Pengantar Pendidikan) then Ya (id = 10)
10. if (nilai == E AND mk_prasyarat == Profesi Pendidikan) then Ya (id = 11)
11. if (nilai == E AND mk_prasyarat == Algoritma dan Pemrograman 2 AND tahun == 2012)
then Ya (id = 13)
12. if (nilai == E AND mk_prasyarat == Algoritma dan Pemrograman 2 AND tahun == 2013)
then Tidak (id = 14)
13. if (nilai == E AND mk_prasyarat == Paket Aplikasi) then Ya (id = 15)
14. if (nilai == E AND mk_prasyarat == Media Pembelajaran Berbasis TIK) then Ya (id = 16)

```

Gambar 7 Tampilan Pohon Keputusan

Pada gambar 4.6 terdapat rule yang mana rule merupakan Id pada setiap atribut yang digunakan, dan dilakukan perhitungan pada database, sehingga membentuk pohon keputusan seperti contoh “if (nilai E == Pengantar Pendidikan) then tidak (id = 10)” yaitu jika matakuliah prasyarat PL pengantar pendidikan maka keputusan ya ID nya 10 keputusan tidak didapat dari

banyak keputusan yang diperoleh yaitu ya, atau tidak.

SIMPULAN

Algoritma C4.5 merupakan model prediksi menggunakan struktur pohon. Pada perancangan program, peneliti hanya memprediksi tahun 2012, 2013 pada mata kuliah syarat PLK di semester ganjil yaitu semester 1, 3, dan 5 yang mana terdiri dari 30 mahasiswa tahun 2012 dan 30 mahasiswa tahun 2013. Simulasi yang dirancang oleh peneliti dalam prediksi mahasiswa mengambil mata kuliah syarat PLK terdiri dari 10 mahasiswa pada 1 kelas.

DAFTAR PUSTAKA

- Alfiansyah, Ilham.2014. Implementasi DataMining Algoritma C4.5 untuk memprediksi jumlah peserta mata kuliah pilihan. Jurnal: STMIK Amikom Yogyakarta (tidak diterbitkan)
- Arief, M.Rudyanto.2011. Pemrograman Web Disain Menggunakan PHP & MySQL. Yogyakarta: Andi.
- A.S.Rosa, dan Shalahuddin,M,2013, Rekayasa Perangkat Lunak Terstruktur dan Berorientasi Objek, Bandung:Informatika.

- Basuki, A dan Syarif,I,2003 Decision Tree, diakses dari <http://www2.eepisits.edu/basuki/lecture/DecisionTree.pdf>, pada 3 maret 2012
- Berry, Micheal J.A and Gordon S.Linoff. 2004 data mining technicquest for marketing sales, customer relationship management. Second edition, wiley publishing, inc
- Fatta, Hanif al.2007. analisis & perancangan sistem informasi. Yogyakarta: andi
- Fayyad, usama.1996. Advance knowledge discovery and data mining. MIT press
- Gafnerz,2013. Dfd, flowchart and erd, diakses dari <http://gafnerz.blogspot.com/2013/06/dfd-flowchart-and-erd.html>
- Glandangan,D. 2012.Rekayasa Perangkat Lunak, diakses dari http://www.academia.edu/7585995/BAB_I_Model_Waterfall
- Ginting, selvia lorena Br, dkk. Analisis dan penerapan algoritma C4.5 dalam data Mining untuk memprediksi masa studi mahasiswa berdasarkan data nilai akademik jurnal. Amikom:bandung
- Kusrini, luthfi,E.T,2009. Algoritma Data Mining. Yogyakarta: Andi.
- Larose, Daniel T.2005. *Discovery Knowledge in Data: An Introduction to Data mining*. John Willey & Sons, Inc
- Lorena,Selvia,dkk.2014.Analisis dan Penerapan Algoritma C4.5 dalam Data Mining untuk Memprediksi Masa Studi Mahasiswa Berdasarkan Data Nilai Akademik. Jurnal: STMIK AMIKOM:Yogyakarta
- Ma'ruf,Faundry Amrul.2013.Aplikasi Data Mining untuk Mengetahui Hubungan Proses Masuk dengan Tingkat Kelulusan Mahasiswa. Jurnal: STMIK AMIKOM:Yogyakarta (Tidak diterbitkan)
- Nugroho,Bunafit.2014. Akademik Sekolah PHP-MySQL Dreamwever. Yogyakarta: Gava Media.
- Ponniah, p. 2001. Datawarehouse Fundamentals: A Comprehensive Guide For IT Professional. John Willey & Sons, Inc.
- Pramudiono, I.2006. Apa Itu data mining? Dalam <http://datamining.japati.net/cgi-bin/indodm.cgi&bacaarsip&1155527614&artikel>. diakses pada tanggal januari 2007
- Suarga, dan Math,M,2012. Algoritma dan Pemrograman. Yogyakarta: Andi.
- Susanto, S. Dan Suryani, D.2010.Pengantar Data Mining.Yogyakarta: Andi.
- Syamsuddin, Aries.2012. Algoritma Decision Tree diakses dari <http://blogs.itb.ac.id/aiceware/2012/09/23/algoritma-decision-tree-c4-5/>
- Tohari, hamim, 2013, ASTAH- analisis serta perancangan sistem informasi melalui pendekatan UML. Yogyakarta : andi
- Turban, E. Dkk 2005 *decision tree support system and intelegent system* yogyakarta : andi offset
- Turban, Usama.1996. Advances in knowledge discovery and data mining.MIT press

