

Teknik Data Mining : Algoritma C 4.5

Rina Novita
noviarinag@gmail.com

Lisensi Dokumen:

Copyright © 2003-2016 IlmuKomputer.Com

Seluruh dokumen di IlmuKomputer.Com dapat digunakan, dimodifikasi dan disebarluaskan secara bebas untuk tujuan bukan komersial (nonprofit), dengan syarat tidak menghapus atau merubah atribut penulis dan pernyataan copyright yang disertakan dalam setiap dokumen. Tidak diperbolehkan melakukan penulisan ulang, kecuali mendapatkan ijin terlebih dahulu dari IlmuKomputer.Com.

Penerapan teknologi informasi yang pesat di berbagai sosial, bisnis, dan pemerintahan telah menghasilkan tumpukan data yang jika dibiarkan begitu saja akan menjadi kuburan data dan tidak memiliki arti sama sekali. Dengan teknologi Data Mining, data yang menggunakan tersebut dapat diolah kembali menjadi pengetahuan yang berguna. Algoritma C4.5 merupakan kelompok algoritma decision tree. Algoritma ini mempunyai input berupa training samples dan samples. Training samples merupakan data contoh yang digunakan untuk membangun sebuah tree yang telah diuji kebenarannya. Sedangkan samples merupakan field-field data yang digunakan sebagai parameter dalam klasifikasi data.

Kata Kunci: Data Mining, Klasifikasi, Algoritma C4.5

Pendahuluan

Perkembangan teknologi informasi yang semakin canggih saat ini, telah menghasilkan banyak tumpukan data. Pertambahan data yang semakin banyak akan menimbulkan pertanyaan besar, yaitu “apa yang dapat dilakukan dari tumpukan data tersebut?”. Untuk menjawab pertanyaan tersebut, dapat diterapkan sebuah teknologi basis data yang dikenal dengan *data mining*.

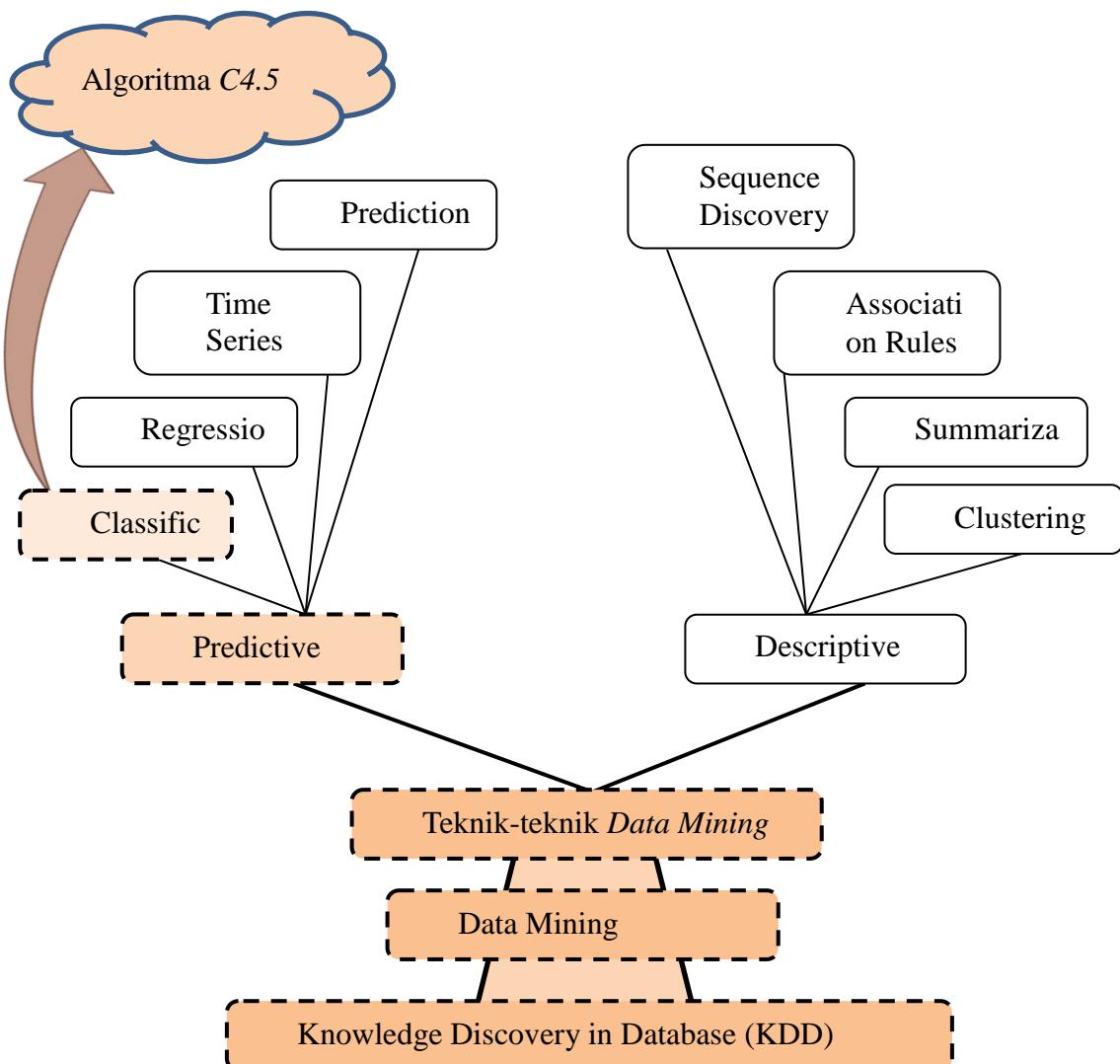
Data mining dapat diterapkan untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual. Terdapat beberapa teknik yang digunakan dalam *data mining*, salah satu teknik *data mining* adalah *classification*. Diantara algoritma yang sering digunakan dalam teknik klasifikasi yaitu algoritma C4.5.

Algoritma C4.5 merupakan kelompok algoritma decision tree. Algoritma ini mempunyai input berupa training samples dan samples. Training samples merupakan data contoh yang digunakan untuk membangun sebuah tree yang telah diuji kebenarannya. Sedangkan samples merupakan field-field data yang digunakan sebagai parameter dalam klasifikasi data.

Pembahasan

1. Peta Konsep/Mind Mapp

Peta konsep di bawah ini bertujuan untuk memudahkan kita dalam memahami materi yang dibahas dalam artikel ini. Peta konsep tentang teknik *data mining* adalah sebagai berikut :

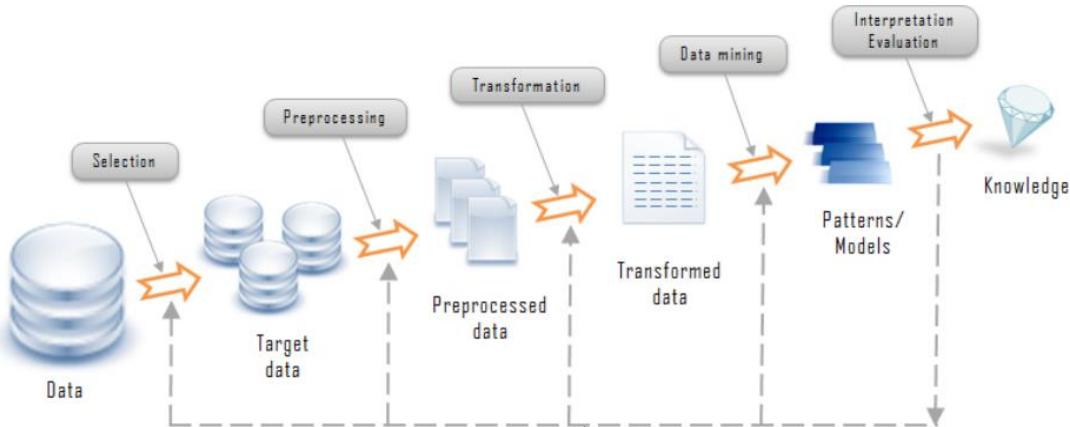


Gambar 1 Peta Konsep/Mind Map Teknik Data Mining

2. Knowledge Discovery in Database (KDD)

Pemrosesan data yang dilakukan secara komputerisasi dalam berbagai bidang, baik dalam penanganan transaksi bisnis maupun untuk kalangan pemerintah dan sosial telah menghasilkan data yang berukuran besar. Data yang dihasilkan tersebut merupakan tambang emas yang dapat diolah untuk mendapatkan pengetahuan. Data yang berskala besar tersebut dapat diolah dengan menggunakan teknologi *data mining* yang merupakan salah satu tahapan dalam *Knowledge Discovery in Database* (KDD).

KDD adalah proses terorganisir untuk mengidentifikasi pola yang valid, baru, berguna, dan dapat dimengerti dari sebuah *data set* yang besar dan kompleks (Maimon dan Lior, 2010). Tahapan dalam KDD terdiri atas pembersihan data (*data cleaning*), integrasi data (*data integration*), pemilihan data (*data selection*), transformasi data (*data transformation*), *data mining*, evaluasi pola (*pattern evaluation*) dan penyajian pengetahuan (*knowledge presentation*) dan penyajian pengetahuan (*knowledge presentation*).



Gambar 2 Proses Knowledge Discovery in Database

Proses KDD secara garis besar dapat dijelaskan sebagai berikut (Angga dan Riani, 2012) :

1. Data Selection

Pemilihan data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam KDD dimulai.

2. Pre-processing/Cleaning

Sebelum proses *data mining* dilaksanakan, perlu dilakukan proses *cleaning* dengan tujuan untuk membuang duplikasi data, memeriksa data yang *inkonsisten*, dan memperbaiki kesalahan pada data, seperti kesalahan cetak. Juga dilakukan proses *enrichment*, yaitu proses “memperkaya” data yang sudah ada dengan data atau informasi lain yang relevan dan diperlukan untuk KDD, seperti data atau informasi *eksternal*.

3. Transformation

Proses *coding* pada data yang telah dipilih, sehingga data tersebut sesuai untuk proses *data mining*. Proses *coding* dalam KDD merupakan proses kreatif dan sangat tergantung pada jenis atau pola informasi yang akan dicari dalam *database*.

4. Data mining

Data mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu.

5. Interpretation/ Evaluation

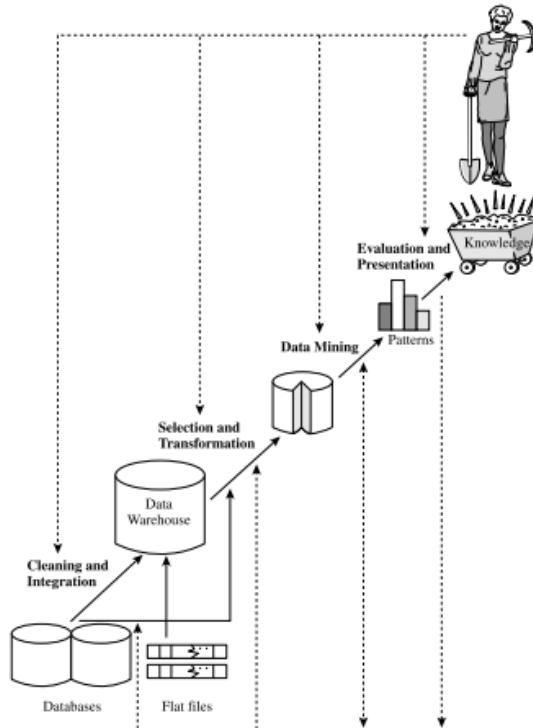
Pola informasi yang dihasilkan dari proses *data mining* perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan. Tahap ini merupakan bagian dari proses KDD yang disebut *interpretation*. Tahap ini mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesis yang ada sebelumnya atau tidak.

3. Data Mining

Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai *database* besar (Mujib, et al, 2013). *Data mining* merupakan salah satu cabang ilmu komputer yang relatif baru. Definisi lain dari *Data mining* dapat diartikan sebagai serangkaian proses untuk menggali nilai tambah berupa informasi yang selama ini tidak diketahui secara manual dari suatu basisdata dengan melakukan proses ekstraksi dan menggali pola penting dari data yang ada (Rizky dan Nita, 2013).

a. Tahapan dalam *Data Mining*

Tuntutan dari *data mining* salah satunya adalah ketika diterapkan pada data berskala besar adalah diperlukan metodologi sistematis tidak hanya ketika melakukan analisa saja tetapi juga ketika mempersiapkan data dan juga melakukan interpretasi dari hasilnya sehingga dapat menjadi aksi ataupun keputusan yang bermanfaat. Sebagai suatu rangkaian proses, *data mining* dapat dibagi menjadi beberapa tahap proses yang diilustrasikan pada gambar 2. Tahap-tahap tersebut bersifat interaktif, pemakai terlibat langsung atau dengan perantaraan *knowledge base* (Mujib, et al, 2008)



Gambar 3 Tahap-tahap *Data Mining*

Tahap-tahap *data mining* adalah sebagai berikut:

1. Pembersihan data (*data cleaning*)

Pembersihan data merupakan proses menghilangkan *noise* dan data yang tidak konsisten atau data tidak relevan.

2. Integrasi data (*data integration*)

Integrasi data merupakan penggabungan data dari berbagai *database* ke dalam satu *database* baru.

3. Seleksi data (*data selection*)

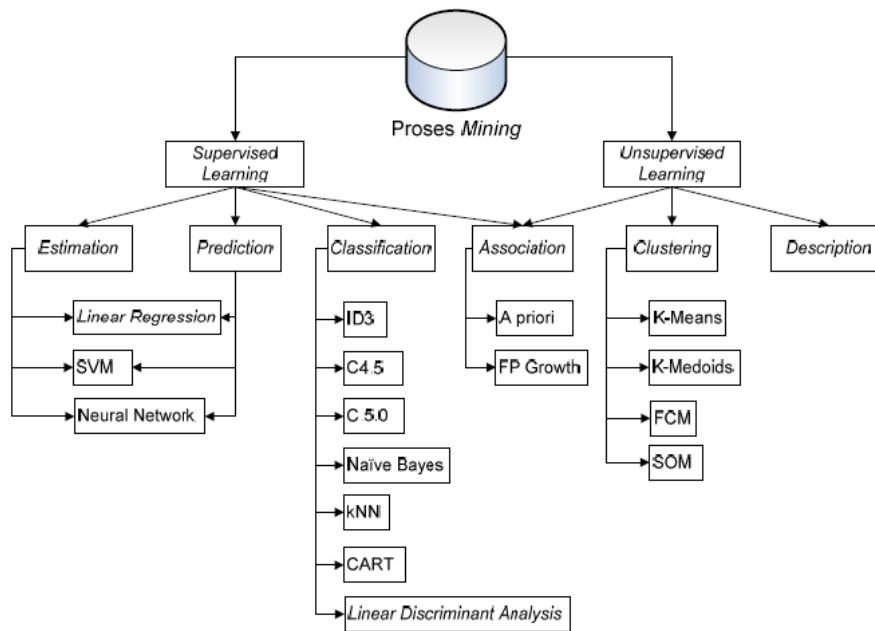
Data yang ada pada *database* sering kali tidak semuanya dipakai, oleh karena itu hanya data yang sesuai untuk dianalisis yang akan diambil dari *database*.

4. Transformasi data (*data transformation*)

Data diubah atau digabung ke dalam format yang sesuai untuk diproses dalam *data mining*.

5. Proses *mining*

Merupakan suatu proses utama saat metode diterapkan untuk menemukan pengetahuan berharga dan tersembunyi dari data. Beberapa metode yang dapat digunakan berdasarkan pengelompokan *data mining* dapat dilihat pada Gambar 2.3



Gambar 4 Beberapa Metode *Data Mining*

6. Evaluasi pola (pattern evaluation)

Untuk mengidentifikasi pola-pola menarik ke dalam *knowledge based* yang ditemukan.

7. Presentasi pengetahuan (*knowledge presentation*)

Merupakan visualisasi dan penyajian pengetahuan mengenai metode yang digunakan untuk memperoleh pengetahuan yang diperoleh pengguna.

b. Pengelompokan *Data mining*

Pengelompokan *data mining* dibagi menjadi beberapa kelompok, menurut Kusrini dan Luthfi (2009), yaitu:

1. Deskripsi

Deskripsi merupakan cara untuk menggambarkan pola dan kecenderungan yang terdapat dalam data yang dimiliki.

2. Estimasi

Estimasi hampir sama dengan klasifikasi, kecuali *variable* target estimasi lebih ke arah numerik daripada kearah kategori. Model yang dibangun menggunakan *record* lengkap yang menyediakan nilai *variable* target sebagai nilai prediksi.

3. Prediksi

Prediksi menerka sebuah nilai yang belum diketahui dan juga memperkirakan nilai untuk masa mendatang.

4. Klasifikasi

Dalam klasifikasi terdapat target variabel kategori, misal penggolongan pendapatan dapat dipisahkan dalam tiga kategori, yaitu tinggi, sedang, dan rendah.

5. Pengklasteran

Merupakan pengelompokan *record*, pengamatan, atau memperhatikan dan membentuk kelas objek-objek yang memiliki kemiripan.

6. Asosiasi

Asosiasi bertugas menemukan atribut yang muncul dalam satu waktu. Dalam dunia bisnis lebih umum disebut analisis keranjang belanja.

4. Klasifikasi

Seiring dengan perkembangan pengetahuan *data mining* dan komponen-komponennya, *data mining* tidak lagi dimonopoli oleh bidang teknologi informasi. Pemakainya telah semakin meluas ke bidang lain misalnya pada bidang kesehatan, pertanian, asuransi, dan lain-lain (Mardiani, 2012).

Khusus untuk fungsionalitas *data mining* menggunakan klasifikasi, beberapa penelitian telah dilakukan misalnya oleh Wisaeng (2013) yang menggunakan pembentukan pohon klasifikasi untuk UCI *repository* dengan membandingkan efektifitas dari beberapa algoritma yaitu algoritma *functional tree*, *logistic model trees*, REP tree, dan best first decision tree. Penelitian Rathee dan Mathur (2013) mengaplikasikan algoritma ID3, algoritma C4.5 dan CART untuk memprediksi kinerja mahasiswa dalam ujian akhir, sehingga membantu pengajar dalam mengidentifikasi kelemahan dan memperbaiki kinerja mereka.

Penelitian-penelitian yang telah dilakukan berkaitan dengan klasifikasi *data mining* juga dilakukan oleh Ogheneovo dan Asagba (2013) terhadap klasifikasi SQL query khususnya terhadap SQL injection attack menggunakan algoritma C4.5. Pinky (2014) dalam mengklasidikasikan lalu lintas menggunakan pendekatan supervised learning dan building time sehingga mampu mendeteksi pengacau lalu lintas dan memberikan pelayanan terbaik kepada users. Penelitian lain dilakukan oleh Patil *et al* (2014) dalam melakukan klasifikasi terhadap data yang tidak pasti (*uncertain*) dengan tujuan mengukur kesalahan untuk berbagai kelas dengan menggunakan algoritma C4.5.

5. Algoritma C4.5

Algoritma yang dapat digunakan untuk memprediksi atau mengklasifikasi suatu kejadian dengan pembentukan pohon keputusan antara lain algoritma C4.5, yang merupakan salah satu algoritma induksi pohon keputusan yang dikembangkan oleh J.Ross Quinlan (Yamin, *et al*, 2014).

Algoritma C4.5 merupakan kelompok algoritma *decision tree*. Algoritma ini mempunyai input berupa *training samples* dan *samples*. *Training samples* merupakan data contoh yang digunakan untuk membangun sebuah *tree* yang telah diuji kebenarannya. Sedangkan *samples* merupakan *field-field* data yang digunakan sebagai parameter dalam klasifikasi data (Sujana, 2010).

Secara umum alur proses algoritma C4.5 untuk membangun pohon keputusan dalam *data mining* adalah sebagai berikut (Mahmud, *et al*, 2014):

- a. Pilih atribut sebagai akar
- b. Buat cabang untuk tiap-tiap nilai
- c. Bagi kasus dalam cabang
- d. Ulangi proses untuk setiap cabang sampai semua kasus pada cabang memiliki kelas yang sama.

Information gain adalah atribut *selection measure* yang digunakan untuk memilih test atribut tiap node dalam *tree* (Sujana, 2010). Pemilihan atribut sebagai simpul, baik simpul akar (*root*) atau simpul internal didasarkan pada nilai *gain* tertinggi dari atribut-atribut yang ada. Perhitungan nilai *gain* digunakan rumus sebagai berikut (Mahmud, *et al*, 2014):

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

Keterangan :

- S : himpunan kasus
- A : atribut
- N : jumlah partisi atribut A
- [S_i] : jumlah kasus pada partisi ke-i
- [S] : jumlah kasus pada S

Untuk menghitung *Entropy* menggunakan rumus :

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

Keterangan :

- S : himpunan kasus
- A : fitur
- N : jumlah partisi S
- p_i : proporsi dari S_i terhadap S

Contoh kasus penggunaan algoritma C4.5 terlihat pada tabel 1 :

Tabel 1 Keputusan Bermain Tenis

NO	OUTLOOK	TEMPERATURE	HUMIDITY	WINDY	PLAY
1	Sunny	Hot	High	FALSE	No
2	Sunny	Hot	High	TRUE	No
3	Cloudy	Hot	High	FALSE	Yes
4	Rainy	Mild	High	FALSE	Yes
5	Rainy	Cool	Normal	FALSE	Yes
6	Rainy	Cool	Normal	TRUE	Yes
7	Cloudy	Cool	Normal	TRUE	Yes
8	Sunny	Mild	High	FALSE	No
9	Sunny	Cool	Normal	FALSE	Yes
10	Rainy	Mild	Normal	FALSE	Yes
11	Sunny	Mild	Normal	TRUE	Yes
12	Cloudy	Mild	High	TRUE	Yes
13	Cloudy	Hot	Normal	FALSE	Yes
14	Rainy	Mild	High	TRUE	No

Pada kasus yang terdapat pada tabel 2.1 akan dibuat pohon keputusan untuk menentukan main tenis atau tidak dengan melihat keadaan cuaca, *temperature*, kelembapan, dan keadaan angin.

Langkah-langkah penyelesaian kasus pada table 2.1 adalah:

- Menghitung jumlah kasus, jumlah kasus untuk keputusan *Yes*, jumlah kasus untuk keputusan *No*, dan *Entropy* dari semua kasus dan kasus yang dibagi berdasarkan atribut *OUTLOOK*, *TEMPERATUR*, *HUMIDITY*, dan *WINDY*.
- Lakukan perhitungan *Gain* untuk setiap atribut. Hasil perhitungan ditunjukkan oleh tabel 2.

Tabel 2 Perhitungan Node 1

Node			Jml Kasus (S)	Tidak (S1)	Ya (S2)	Entropy	Gain
1	TOTAL		14	4	10	0.8631205	
	OUTLOOK						0.2585210
		Cloudy	4	0	4		
		Rainy	5	1	4	0.721928	
		Sunny	5	3	2	0.970950	
	TEMPERATURE						0.1838509
		Cool	4	0	4	0	
		Hot	4	2	2	1	
		Mild	6	2	4	0.9182958	
	HUMADITY						0.3705065
		High	7	4	3	0.9852281	
		Normal	7	0	7	0	
	WINDY						0.0059777
		False	8	2	6	0.8112781	
		True	6	4	2	0.9182958	

Baris total kolom *Entropy* pada table 2 dihitung dengan persamaan berikut:

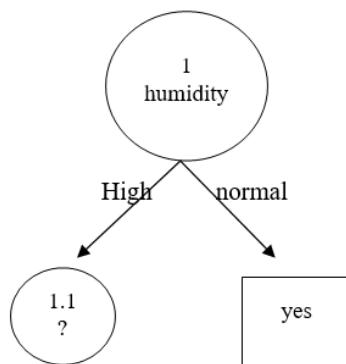
$$\begin{aligned} \text{Entropy (Total)} &= \left(-\frac{4}{14} * \log_2 \left(\frac{4}{14} \right) \right) + \left(-\frac{10}{14} * \log_2 \left(\frac{10}{14} \right) \right) \\ &= 0.863120569 \end{aligned}$$

Nilai *Gain* pada baris *OUTLOOK* dihitung dengan persamaan berikut:

$$\begin{aligned} \text{Gain}(Total, Outlook) &= \text{Entropy}(Total) - \sum_{i=1}^n \frac{|Outlook|}{|Total|} * \text{Entropy}(Outlook) \\ &= 0.863120569 - \left(\left(\frac{4}{14} * 0 \right) + \left(\frac{5}{14} * 0.723 \right) + \left(\frac{5}{14} * 0.97 \right) \right) \\ &= 0.23 \end{aligned}$$

Dari tabel 2.2 dapat diketahui bahwa atribut dengan *Gain* tertinggi adalah *HUMADITY*, yaitu sebesar 0.37. Jadi *HUMADITY* yang menjadi *node* akar. Ada dua nilai atribut dari *HUMADITY*, yaitu *HIGH* dan *NORMAL*. Dari kedua atribut, nilai atribut *NORMAL* adalah 1, yaitu keputusan *Yes*, sehingga tidak perlu perhitungan lebih lanjut, tetapi untuk *HIGH* masih perlu dilakukan perhitungan lagi.

Dari hasil tersebut dapat digambarkan pohon keputusan sementara seperti gambar 5 berikut,



Gambar 5. Pohon Keputusan Haisl Perhitungan Note 1

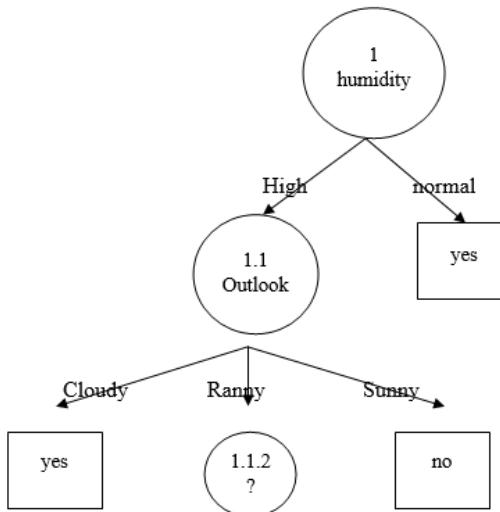
Menghitung jumlah kasus, jumlah kasus untuk keputusan *Yes*, jumlah kasus untuk keputusan *No*, dan *Entropy* dari semua kasus dan kasus yang dibagi berdasarkan atribut *OUTLOOK*, *TEMPERATURE*, dan *WINDY* yang dapat menjadi node akar dari nilai atribut *HIGH*. Kemudian lakukan perhitungan *Gain* untuk tiap-tiap atribut. Hasil perhitungan seperti yang terdapat pada tabel 3 :

Tabel 3 Perhitungan Node 1.1

Node		Jml Kasus (S)	Tidak (S1)	Ya (S2)	Entropy	Gain
1.1	TOTAL	7	4	3	0.9852281	
	OUTLOOK					0.69951385
	Cloudy	2	0	2	0	
	Rainy	2	1	1	1	
	Sunny	3	3	0	0	
	TEMPERATURE					0.02024420
	Cool	0	0	0	0	
	Hot	4	2	1	0.9182958	
	Mild	3	2	2	1	
	WINDY					0.02024420
	False	4	2	2	1	
	True	3	2	1	0.9182958	

Dari hasil tabel 3 dapat diketahui bahwa atribut dengan *Gain* tertinggi adalah *OUTLOOK*, yaitu sebesar 0.67. Jadi *OUTLOOK* dapat menjadi node cabang dari nilai atribut *HIGH*. Ada tiga nilai atribut dari *OUTLOOK*, yaitu *CLOUDY*, *RAINY* dan *SUNNY*. Dari ketiga nilai atribut tersebut, nilai atribut *CLOUDY* adalah 1, yaitu keputusannya *Yes* dan nilai atribut *SUNNY* menjadi keputusan *No*, sehingga tidak perlu dilakukan perhitungan lebih lanjut, tetapi nilai atribut *RAINY* masih perlu dilakukan perhitungan lagi.

Pohon keputusan yang terbentuk sampai tahap ini adalah terlihat pada gambar 6 berikut.



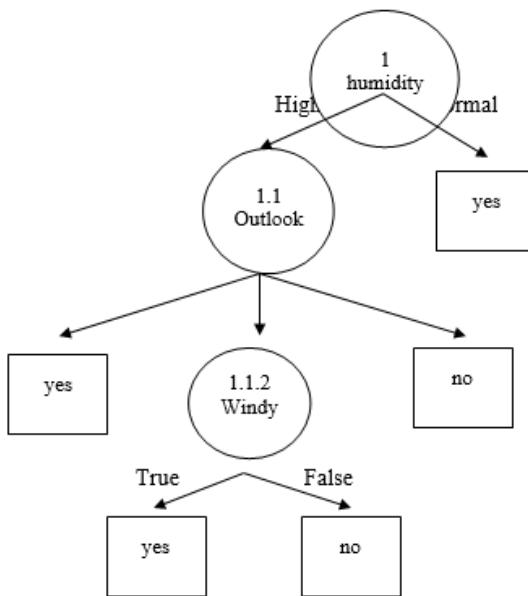
Gambar 6. Pohon Keputusan Hasil Perhitungan Node 1.1

Menghitung jumlah kasus, jumlah kasus untuk keputusan Yes, jumlah kasus untuk keputusan No, dan Entropy dari semua kasus dan kasus yang dibagi berdasarkan atribut *TEMPERATURE* dan *WINDY* yang dapat menjadi *node* cabang dari nilai atribut *RAINY*. Kemudian lakukan perhitungan *Gain* untuk tiap-tiap atribut. Hasil perhitungan seperti yang terdapat pada tabel 4 :

Tabel 4 Perhitungan Node 1.1.2

Node			Jml Kasus (S)	Tidak (S1)	Ya (S2)	Entropy	Gain
1.1.2	HUMADITY-HIGH dan OUTLOOK-RAINY		2	1	1	1	
	TEMPERATURE					0	
		Cool	0	0	0	0	
		Hot	0	0	0	0	
		Mild	2	1	1	1	
	WINDY					1	
		False	1	0	1	0	
		True	1	1	0	0	

Dari hasil tabel 4 dapat diketahui bahwa atribut dengan *Gain* tertinggi adalah *WINDY*, yaitu sebesar 1. Jadi *WINDY* dapat menjadi *node* cabang dari nilai atribut *RAINY*. Ada dua nilai atribut dari *WINDY*, yaitu *FALSE* dan *TRUE*. Dari kedua nilai atribut tersebut, nilai atribut *FALSE* adalah 1, yaitu keputusannya Yes dan nilai atribut *TRUE* menjadi keputusan No, sehingga tidak perlu dilakukan perhitungan lebih lanjut untuk nilai atribut ini. Pohon keputusan yang terbentuk sampai tahap ini adalah terlihat pada gambar 7 berikut:



Gambar 7 Pohon Keputusan Hasil Perhitungan *Node 1.1.2*

Gambar 7 merupakan gambar akhir dari proses *Data Mining* menggunakan algoritma C4.5 pada contoh kasus keputusan bermain tenis. Berdasarkan gambar di atas dapat ditarik kesimpulan bahwa jika *humadity* normal maka keputusannya adalah *yes*, jika *humadity* adalah *high*, *outlook* adalah *cloudy* maka keputusannya adalah bermain (*yes*), selanjutnya jika *humadity* adalah *high*, *outlook* adalah *rainy* dan *windy* adalah *true*, maka keputusannya adalah bermain (*yes*). Dan jika *humadity* adalah *high*, *outlook* adalah *rainy* dan *windy* adalah *false*, maka keputusannya adalah tidak bermain (*no*). Terakhir jika *humadity* adalah *high*, *outlook* adalah *no*, maka keputusannya adalah tidak bermain (*no*).

Penutup

Algoritma C4.5 merupakan kelompok algoritma decision tree. Algoritma ini mempunyai input berupa training samples dan samples. Training samples merupakan data contoh yang digunakan untuk membangun sebuah tree yang telah diuji kebenarannya. Sedangkan samples merupakan field-field data yang digunakan sebagai parameter dalam klasifikasi data.

Referensi

- MABRUR, A. G. & LUBIS, R. 2012. Penerapan Data Mining untuk Memprediksi Kriteria Nasabah Kredit. Int J Cancer.
- MAIMON, O. & ROKASH, L. 2010. Data Mining and Knowledge Discovery Handbook.
- MARDIANI 2012. Penerapan Klasifikasi dengan Algoritma Cart untuk Prediksi Kuliah Bagi Mahasiswa Baru. Seminar Nasional Aplikasi Teknologi Informasi, 13, 584.
- OGHENEOVO, E. E. & ASAGBA, P. O. 2013. Application of a Modified C4.5 Learner's Algorithm to SQL Query Classification. African Journal of Computing & ICT, 6, 37-48.
- PATIL, K., PAGARE, N., NARHEDE, P. & BRAHMANKAR, P. 2014. Classifying Climate Data (uncertain) Using Decision Tree. International Journal of Advanced Research, 2, 402-408.
- PINKY, P. 2014. Clssification of Traffic uing Supervised Learning Approach and its Building Time. The International Journal of Science & Technoledge, 2, 91-93.
- RATHEE, A. & MATHUR, R. P. 2013. Survey on Decision Three Classification Algorithms for Evaluation of Student Performance. International Journal of Computers & Technology, 4 4.
- RIDWAN, M., SUYONO, H. & SAROSA, M. 2013. Penerapan Data Mining untuk Evaluasi Kinerja

Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier. *EECCIS*, 7, 858-71. SUNJANA 2010. Klasifikasi Data Nasabah Sebuah Asuransi Menggunakan Algoritma C4.5. SNATI, 31-34.

WISAENG, K. 2013. A Comparison of Decision Tree Algorithms for UCI Repository Classification. *International Journal of Engineering Trends and Technology (IJETT)*, 4, 5.

YUNUS, M., DAHLAN, H. S. & SANTOSO, P. B. 2014. SPK Pemilihan Calon Pendonor Darah Potensial dengan Algoritma C4.5 dan Fuzzy Tahani. *EECCIS*, 8.

Biografi Penulis



Rina Novita. Lahir di Parik Lintang, 06 November 1990, Menyelesaikan S1 di Prodi Pendidikan Teknik Informatika dan Komputer STAIN Bukittinggi pada April 2013, dan menyelesaikan S2 di Universitas Putra Indonesia YPTK Padang pada Oktober 2014. Sekarang menjadi Dosen pada Prodi Pendidikan Teknik Informatika dan Komputer Institut Agama Islam Negeri Bukittinggi, Sumatera Barat, Indonesia.