**BAB III**

**ANALISA DAN PERANCANGAN**

**3.1 Analisa Permasalahan Metode Decision Tree**

Ada beberapa tahapan dalam membuat sebuah *decision tree* (pohon keputusan) dalam algoritma C4.5 yaitu :

1. Mempersiapkan data *training*

Data *training* biasanya diambil dari data riwayat yang pernah terjadi sebelumnya atau disebut masa lalu dan sudah dikelompokkan dalam kelas-kelas tertentu

1. Pilih atribut sebagai akar

Untuk memilih atribut sebagai akar, dengan cara menghitung nilai *gain* dari masing-masing atribut, nilai *gain* yang paling tinggi akan menjadi akar pertama. Sebelum menghitung nilai *gain* dari atribut, hitung dahulu nilai *entropy*. Untuk menghitung nilai *entropy* didapatkan rumus sebagai berikut :

*Gain (S,A) Entropy (S) .................* (1)

di mana nilai *Entropy* didapatkan dengan rumus :

*Entropy (S) = ....................* (2)

dan nilai p didapatkan rumus :

*p = ...............................*............................................ (3)

sehingga didapatkan nilai *Entropy* sebagai berikut:

*Entropy (S) = .............................................* (4)

Keterangan :

*S* = Ruang (data) sampel yang digunakan untuk pelatihan

*n* = Jumlah partisi atribut S

*A =*Atribut

= Proporsi ke-*i*

|Si| = Jumlah kasus ke-*i*

|S| = Jumlah kasus ke *S*

= Jumlah yang bersolusi positif atau mendukung pada data sampel kriteria tertentu

= Jumlah yang bersolusi negatif atau tidak mendukung pada data sampel kriteria tertentu.

1. Buat cabang masing-masing nilai
2. Bagi kasus dalam cabang
3. Ulangi proses untuk masing-masing cabang sampai semua kasus pada cabang memiliki kelas yang sama.

Dalam kasus ini akan diambil dari data prakiraan cuaca untuk kota–kota di Kabupaten/Kota di Sumatera Utara yang ada di Kantor BMKG Wil 1 Medan di tanggal 15 April 2016. Data tersebut akan digunakan dalam bentuk pohon keputusan untuk menganalisa prakiraan cuaca berawan, hujan atau cerah. Data tersebut selanjutnya akan dilakukan pra-proses untuk menghasilkan data khusus yang siap untuk dibentuk menjadi sebuah pohon keputusan.

Tabel 3.1 Data Prakiraan Cuaca

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **No** | **Kota/Kab** | **Suhu max** | **Suhu min** | **Kelembaban** | **Arah angin** | **Cuaca** |
|
| 1 | Pandan | 33 | 22 | 79 | Barat Daya/10 KT | Berawan |
| 2 | Tarutung | 31 | 18 | 86 | Barat Daya/10 KT | Hujan |
| 3 | Sipirok | 33 | 23 | 84 | Barat Daya/10 KT | Cerah |
| 4 | Gunung Sitoli | 33 | 23 | 85 | Barat Daya/10 KT | Cerah |
| 5 | Stabat | 34 | 24 | 85 | Barat Daya/10 KT | Cerah |
| 6 | Kabanjahe | 29 | 17 | 80 | Barat Daya/10 KT | Hujan |
| 7 | Lubuk Pakam | 34 | 24 | 80 | Barat Daya/10 KT | Cerah |
| 8 | Pematang Raya | 31 | 18 | 80 | Barat Daya/10 KT | Hujan |
| 9 | Kisaran | 34 | 23 | 80 | Barat Daya/10 KT | Cerah |
| 10 | Rantau Prapat | 34 | 23 | 80 | Barat Daya/10 KT | Berawan |
| 11 | Sidikalang | 29 | 18 | 80 | Barat Daya/10 KT | Hujan |
| 12 | Balige | 30 | 17 | 80 | Barat Daya/10 KT | Hujan |
| 13 | Panyabungan | 33 | 23 | 80 | Barat Daya/10 KT | Berawan |
| 14 | Teluk Dalam | 33 | 23 | 82 | Barat Daya/10 KT | Cerah |
| 15 | Salak | 30 | 18 | 85 | Barat Daya/10 KT | Hujan |
| 16 | Dolok Sanggul | 29 | 17 | 87 | Barat Daya/10 KT | Hujan |
| 17 | Pangururan | 29 | 17 | 87 | Barat Daya/10 KT | Hujan |
| 18 | Sei Rampah | 34 | 24 | 75 | Barat Daya/10 KT | Cerah |

Tabel 3.1 Data Prakiraan Cuaca (lanjutan)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **No** | **Kota/Kab** | **Suhu max** | **Suhu min** | **Kelembaban** | **Arah angin** | **Cuaca** |
|
| 19 | Lima Puluh | 34 | 24 | 78 | Barat Daya/10 KT | Cerah |
| 20 | Medan | 35 | 24 | 76 | Barat Daya/10 KT | Cerah |
| 21 | Pematang Siantar | 32 | 21 | 82 | Barat Daya/10 KT | Hujan |
| 22 | Sibolga | 33 | 21 | 78 | Barat Daya/10 KT | Berawan |
| 23 | Tanjung Balai | 34 | 24 | 85 | Barat Daya/10 KT | Cerah |
| 24 | Binjai Kota | 34 | 24 | 75 | Barat Daya/10 KT | Cerah |
| 25 | Tebing Tinggi | 33 | 23 | 76 | Barat Daya/10 KT | Cerah |
| 26 | Padang Sidempuan | 34 | 23 | 75 | Barat Daya/10 KT | Berawan |
| 27 | Kota pinang | 34 | 23 | 75 | Barat Daya/10 KT | Berawan |
| 28 | Aek Kanopan | 33 | 23 | 75 | Barat Daya/10 KT | Berawan |
| 29 | Lahomi | 33 | 23 | 75 | Barat Daya/10 KT | Cerah |
| 30 | Lotu | 33 | 23 | 75 | Barat Daya/10 KT | Cerah |
| 31 | Sibuhan | 31 | 23 | 80 | Barat Daya/10 KT | Hujan |
| 32 | Gunung Tua | 31 | 23 | 80 | Barat Daya/10 KT | Hujan |

*Sumber data : BMKG Wil 1 15April 2016*

Data pada tabel 3.1 sangat banyak dan tidak konsisten pada umumnya terjadi di setiap *database*. Data yang tidak lengkap disebabkan karena adanya data yang belum normal atau atribut yang salah, sehingga proses data *preprocessing* perlu dilakukan sehingga *database* sesuai dengan ketentuan yang diperlukan.

**3.2 Algoritma Sistem**

Algoritma merupakan langkah-langkah dalam penyelesaian suatu masalah, dalam penyelesaian metode *decision tree* algoritma C4.5, maka dilakukan tahapan-tahapan sebagai berikut :

1. Data *selection*
2. Data *preprocessing*/*cleaning*
3. Mengubah data menjadi *tree*
4. Menentukan *node* terpilih
5. Mengubah *tree* menjadi *rule*

**3.2.1 Data Selection**

Data prakiraan cuaca untuk kota–kota di Kabupaten/Kota di Sumatera Utara menjadi data kasus dalam proses operasional *data mining*. Dari data yang ada, kolom yang diambil sebagai atribut keputusan adalah cuaca, sedangkan kolom yang diambil penentuan dalam pembentukan pohon keputusan adalah :

1. Kabupaten/Kota

Merupakan nama Kabupaten/Kota di Sumatera utara dalam memprediksi prakiraan cuaca

1. Suhu Maksimum

Suhu maksimum di bagi 2 kategori yaitu :

1. Di bawah 30 : jika suhu ≤ 30
2. Di atas 30 : jika suhu > 30
3. Suhu Minimum

Suhu minimum di bagi 2 kategori yaitu :

1. Di bawah 23 : jika suhu ≤ 23
2. Di atas 23 : jika suhu > 23
3. Kelembaban

Kelembaban di bagi 2 kategori yaitu :

1. Di bawah 80 : jika suhu ≤ 80
2. Di atas 80 : jika suhu > 80
3. Arah angin

Arah angin yang terjadi adalah Barat Daya/10 KT.

**3.2.2 Data Prepocessing/Data Cleaning**

*Data cleaning* diterapkan untuk menambah isi atribut yang hilang atau kosong, dan merubah data yang tidak konsisten.

1. Data Transformation

Dalam proses ini, data ditransformasikan ke dalam bentuk yang sesuai untuk proses *data mining.*

1. Data Reduction

Reduksi data dilakukan dengan menghilangkan atribut yang tidak perlukan sehingga ukuran dari *database* menjadi kecil dan hanya menyertakan atribut yang diperlukan dalam proses *data mining,* karena akan lebih efisien terhadap data yang lebih kecil.

Masalah klasifikasi berakhir dengan dihasilkan sebuah pengetahuan yang dipresentasikan dalam bentuk diagram yang biasa disebut pohon keputusan (*decision tree*) untuk menentukan prakiraan cuaca berawan, hujan atau cerah, kriteria yang diperhatikan adalah suhu maksimum, suhu minimum, kelembaban, dan arah angin tersebut.

Berikut ini data yang dipergunakan untuk menentukan berawan, hujan atau cerah dalam prakiraan cuaca, data selengkapnya pada tabel 3.2 berikut ini :

Tabel 3.2 Hasil Normalisasi Data

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **No** | **Kota/Kab** | **Suhu max** | **Suhu min** | **Kelembaban** | **Arah angin** | **cuaca** |
|
| 1 | Pandan | Di atas 30 | Di bawah 23 | Di bawah 80 | Barat Daya/10 KT | Berawan |
| 2 | Tarutung | Di atas 30 | Di bawah 23 | Di atas 80 | Barat Daya/10 KT | Hujan |
| 3 | Sipirok | Di atas 30 | Di bawah 23 | Di atas 80 | Barat Daya/10 KT | Cerah |
| 4 | Gunung Sitoli | Di atas 30 | Di bawah 23 | Di atas 80 | Barat Daya/10 KT | Cerah |
| 5 | Stabat | Di atas 30 | Di atas 23 | Di atas 80 | Barat Daya/10 KT | Cerah |
| 6 | Kabanjahe | Di bawah 30 | Di bawah 23 | Di bawah 80 | Barat Daya/10 KT | Hujan |
| 7 | Lubuk Pakam | Di atas 30 | Di atas 23 | Di bawah 80 | Barat Daya/10 KT | Cerah |
| 8 | Pematang Raya | Di atas 30 | Di bawah 23 | Di bawah 80 | Barat Daya/10 KT | Hujan |
| 9 | Kisaran | Di atas 30 | Di bawah 23 | Di bawah 80 | Barat Daya/10 KT | Cerah |
| 10 | Rantau Prapat | Di atas 30 | Di bawah 23 | Di bawah 80 | Barat Daya/10 KT | Berawan |
| 11 | Sidikalang | Di bawah 30 | Di bawah 23 | Di bawah 80 | Barat Daya/10 KT | Hujan |
| 12 | Balige | Di bawah 30 | Di bawah 23 | Di bawah 80 | Barat Daya/10 KT | Hujan |
| 13 | Panyabungan | Di atas 30 | Di bawah 23 | Di bawah 80 | Barat Daya/10 KT | Berawan |

Tabel 3.2 Hasil Normalisasi Data (lanjutan)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **No** | **Kota/Kab** | **Suhu max** | **Suhu min** | **Kelembaban** | **Arah angin** | **cuaca** |
|
| 14 | Teluk Dalam | Di atas 30 | Di bawah 23 | Di atas 80 | Barat Daya/10 KT | Cerah |
| 15 | Salak | Di bawah 30 | Di bawah 23 | Di atas 80 | Barat Daya/10 KT | Hujan |
| 16 | Dolok Sanggul | Di bawah 30 | Di bawah 23 | Di atas 80 | Barat Daya/10 KT | Hujan |
| 17 | Pangururan | Di bawah 30 | Di bawah 23 | Di atas 80 | Barat Daya/10 KT | Hujan |
| 18 | Sei Rampah | Di atas 30 | Di atas 23 | Di bawah 80 | Barat Daya/10 KT | Cerah |
| 19 | Lima Puluh | Di atas 30 | Di atas 23 | Di bawah 80 | Barat Daya/10 KT | Cerah |
| 20 | Medan | Di atas 30 | Di atas 23 | Di bawah 80 | Barat Daya/10 KT | Cerah |
| 21 | Pematang Siantar | Di atas 30 | Di bawah 23 | Di atas 80 | Barat Daya/10 KT | Hujan |
| 22 | Sibolga | Di atas 30 | Di bawah 23 | Di bawah 80 | Barat Daya/10 KT | Berawan |
| 23 | Tanjung Balai | Di atas 30 | Di atas 23 | Di atas 80 | Barat Daya/10 KT | Cerah |
| 24 | Binjai Kota | Di atas 30 | Di atas 23 | Di bawah 80 | Barat Daya/10 KT | Cerah |
| 25 | Tebing Tinggi | Di atas 30 | Di bawah 23 | Di bawah 80 | Barat Daya/10 KT | Cerah |
| 26 | Padang Sidempuan | Di atas 30 | Di bawah 23 | Di bawah 80 | Barat Daya/10 KT | Berawan |
| 27 | Kota pinang | Di atas 30 | Di bawah 23 | Di bawah 80 | Barat Daya/10 KT | Berawan |

Tabel 3.2 Hasil Normalisasi Data (lanjutan)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **No** | **Kota/Kab** | **Suhu max** | **Suhu min** | **Kelembaban** | **Arah angin** | **cuaca** |
|
| 28 | Aek Kanopan | Di atas 30 | Di bawah 23 | Di bawah 80 | Barat Daya/10 KT | Berawan |
| 29 | Lahomi | Di atas 30 | Di bawah 23 | Di bawah 80 | Barat Daya/10 KT | Cerah |
| 30 | Lotu | Di atas 30 | Di bawah 23 | Di bawah 80 | Barat Daya/10 KT | Cerah |
| 31 | Sibuhan | Di atas 30 | Di bawah 23 | Di bawah 80 | Barat Daya/10 KT | Hujan |
| 32 | Gunung Tua | Di atas 30 | Di bawah 23 | Di bawah 80 | Barat Daya/10 KT | Hujan |

* + 1. **Mengubah Data Menjad Tree**

Dalam mengubah data menjadi *tree* terlebih dahulu data dinyatakan dalam bentuk tabel dengan atribut dan *record*. Atribut menyatakan suatu parameter yang dibuat sebagai kriteria dalam pembentukan *tree*.

* + 1. **Menentukan Node Terpilih/Menentukan Nilai Atribut**

Dalam data sampel tentukan *node* terpilih, yaitu dengan menghitung nilai informasi *gain* masing-masing atribut untuk menentukan *node* terpilih, gunakan nilai informasi *gain* yang paling besar. Sebelum menghitung nilai *gain* dari atribut, hitung dahulu nilai *entropy*.

Menghitung nilai *entropy* dari masing-masing atribut adalah sebagai berikut:

1. Dimulai dari *node* akar, harus dihitung terlebih dahulu *entropy* untuk *node* akar (semua data) terhadap komposisi kelas.

*Entropy (S) =*

*E(S) = -*

*E(S)* ***=***

=

= - (0.3437) ((log (11/32)/log 2) (0.2187) ((log (7/32)/log 2) (0.4375) ((log (14/32)/log 2)

= - (0.3437) (-0.4637/0.3010) – (0.2187) (-0.6600/0.3010) – (0.4375)

(-0.3590/0.3010)

= - (0.3437) (-1.5405) – (0.2187) (-2.1926) – (0.4375) (-1.1926)

= 0.5294 + 0.4795 + 0.5217

= 1.5309

1. Selanjutnya dihitung *entropy* untuk setiap nilai fitur terhadap kelas. Untuk *entropy* nilai dalam atribut sebagai berikut :
   1. *Entropy* suhu maksimum

*E(S,di bawah 30)*

*=*

=

= 0

*E(S,di atas 30)*

=

=

= 1.4479

* 1. *Entropy* suhu minimum

*E(S,di bawah 23)*

*=* =

= 1.5495

*E(S,di atas 23)*

*=*

=

= 0

* 1. *Entropy* kelembaban

*E(S,di bawah 80)*

*=*

=

= 1.5389

*E(S,di atas 80)*

*=*

=

= 0

* 1. *Entropy* arah angin

*E(S,barat daya)*

*=* =

= 1.5309

*E(S,selain barat daya)*

=

= 0

1. Menghitung nilai *Information Gain*

Menghitung nilai *gain* masing-masing dari tiap atribut adalah sebagai berikut :

* 1. Menghitung nilai *gain* atribut suhu maksimum

*G(Semua, suhu max)* = *E*(s)

= 1.5309

= 1.5309 – ( 6/32 \* 0 + 26/32 \* 1.4479)

= 1.5306 – 1.1764

= 0.3545

* 1. Menghitung nilai *gain* atribut suhu minimum

*G(Semua, suhu min)* = *E*(s)

= 1.5309

= 0.3203

* 1. Menghitung nilai *gain* atribut kelembaban

*G(Semua, kelembaban)* = *E*(s)

= 1.5309

= 0.4729

* 1. Menghitung nilai *gain* atribut arah angin

*G(Semua, arah)* = *E*(s)

= 1.5309

= 0

Tabel 3.3 Hasil Perhitungan Entropy dan Gain untuk Node Akar

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Node** |  |  | **Jml Kasus (S)** | Hujan (S1) | Berawan  (S2) | **cerah (S3)** | **Entropy** | **Gain** |
| 1 | Total |  | 32 | 11 | 7 | 14 | 1.5309 |  |
|  | Suhu max |  |  |  |  |  |  | 0.3545 |
|  |  | di bawah 30 | 6 | 6 | 0 | 0 | 0 |  |
|  |  | Di atas 30 | 26 | 5 | 7 | 14 | 1.4479 |  |
|  | Suhu min |  |  |  |  |  |  | 0.3203 |
|  |  | di bawah 23 | 25 | 11 | 7 | 7 | 1.5495 |  |
|  |  | di atas 23 | 7 | 0 | 0 | 7 | 0 |  |
|  | kelembaban |  |  |  |  |  |  | 0.4729 |
|  |  | di bawah 80 | 22 | 5 | 7 | 9 | 1.5389 |  |
|  |  | di atas 80 | 10 | 5 | 0 | 5 | 0 |  |
| arah angin |  |  |  |  |  |  | 0 |
|  | barat daya | 32 | 11 | 7 | 14 | 1.5309 |  |
|  | selain barat daya | 0 | 0 | 0 | 0 | 0 |  |

Setelah informasi selesai dihitung, maka diambil nilai informasi *gain* terbesar dari empat atribut yang ada yaitu atribut kelembaban sebesar 0.4729, jadi kelembaban dapat menjadi *node*/akar awal.

di atas 80 di bawah 80

**Berawan**

Gambar 3.2 Pohon Keputusan Node 1

Pohon keputusan di atas belum terlihat keputusan yang dominan dari setiap cuaca yang dipilih. Maka mencari kembali nilai *entropy* dan *gain* dari setiap atribut kelembaban adalah di bawah 80 sebagai berikut :

1. Menghitung nilai *entropy* total kelembaban (di bawah 80)

Berikut ini adalah tabel penyelesaiannya :

Tabel 3.4 Data Atribut Kelembaban (di bawah 80)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **No** | **Kota/Kab** | **Suhu max** | **Suhu min** | **Arah angin** | **cuaca** |
|
| 1 | Pandan | Di atas 30 | Di bawah 23 | Barat Daya/10 KT | Berawan |
| 2 | Kabanjahe | Di bawah 30 | Di bawah 23 | Barat Daya/10 KT | Hujan |
| 3 | Lubuk Pakam | Di atas 30 | Di atas 23 | Barat Daya/10KT | Cerah |

Tabel 3.4 Data Atribut Kelembaban (di bawah 80) lanjutan

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **No** | **Kota/Kab** | **Suhu max** | **Suhu min** | **Arah angin** | **Cuaca** |
|
| 4 | Pematang Raya | Di atas 30 | Di bawah 23 | Barat Daya/10 KT | Hujan |
| 5 | Kisaran | Di atas 30 | Di bawah 23 | Barat Daya/10 KT | Cerah |
| 6 | Rantau Prapat | Di atas 30 | Di bawah 23 | Barat Daya/10 KT | Berawan |
| 7 | Sidikalang | Di bawah 30 | Di bawah 23 | Barat Daya/10 KT | Hujan |
| 8 | Balige | Di bawah 30 | Di bawah 23 | Barat Daya/10 KT | Hujan |
| 9 | Panyabungan | Di atas 30 | Di bawah 23 | Barat Daya/10 KT | Berawan |
| 10 | Sei Rampah | Di atas 30 | Di atas 23 | Barat Daya/10 KT | Cerah |
| 11 | Lima Puluh | Di atas 30 | Di atas 23 | Barat Daya/10 KT | Cerah |
| 12 | Medan | Di atas 30 | Di atas 23 | Barat Daya/10 KT | Cerah |
| 13 | Sibolga | Di atas 30 | Di bawah 23 | Barat Daya/10 KT | Berawan |
| 14 | Binjai Kota | Di atas 30 | Di atas 23 | Barat Daya/10 KT | Cerah |
| 15 | Tebing Tinggi | Di atas 30 | Di bawah 23 | Barat Daya/10 KT | Cerah |
| 16 | Padang Sidempuan | Di atas 30 | Di bawah 23 | Barat Daya/10 KT | Berawan |
| 17 | Kota pinang | Di atas 30 | Di bawah 23 | Barat Daya/10 KT | Berawan |
| 18 | Aek Kanopan | Di atas 30 | Di bawah 23 | Barat Daya/10 KT | Berawan |
| 19 | Lahomi | Di atas 30 | Di bawah 23 | Barat Daya/10 KT | Cerah |
| 20 | Lotu | Di atas 30 | Di bawah 23 | Barat Daya/10 KT | Cerah |
| 21 | Sibuhan | Di atas 30 | Di bawah 23 | Barat Daya/10 KT | Hujan |
| 22 | Gunung Tua | Di atas 30 | Di bawah 23 | Barat Daya/10 KT | Hujan |

*E(S) =*

=

= 1.5389

1. Menghitung *entropy* untuk setiap nilai fitur terhadap kelas. Untuk *entropy* nilai dalam atribut sebagai berikut :
   1. *Entropy* suhu maksimum

*E(S,di bawah 30)*

*=*

=

= 0

*E(S,di atas 30)*

*=*

=

= 1.4618

* 1. *Entropy* suhu minimum

*E(S,di bawah 23)*

*=*

=

= 1.5485

*E(S,di atas 23)*

*=*

=

= 0

* 1. *Entropy* arah angin

*E(S,barat daya) =*

=

= 1.5389

*E(S,selain barat daya)*

=

= 0

1. Menghitung nilai *Information Gain*

Menghitung nilai *gain* masing-masing dari tiap atribut adalah sebagai berikut :

1. Menghitung nilai *gain* atribut suhu maximum

*G(Semua, suhu max)* = *E*(s)

= 1.5389

= 0.2764

1. Menghitung nilai *gain* atribut suhu minimum

*G(Semua, suhu min)* = *E*(s)  *)*

=1.5389

= 0.3423

1. Menghitung nilai *gain* atribut arah angin

*G(Semua, arah)* = *E*(s)

*=*1.538

= 0

Tabel 3.5 Hasil Perhitungan Entropy dan Gain untuk Node 2

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Node** |  |  | **Jml Kasus (S)** | **Hujan (S1)** | **Berawan(S2)** | **cerah (S3)** | **Entropy** | **Gain** |
| 2 | Kelembaban di bawah 80 |  | 22 | 5 | 7 | 9 | 1.5389 |  |
|  | Suhu max |  |  |  |  |  |  | 0.2764 |
|  | di bawah 30 | 3 | 3 | 0 | 0 | 0 |  |
|  | Di atas 30 | 19 | 3 | 7 | 9 | 1.4618 |  |
| Suhu min |  |  |  |  |  |  | 0.3423 |
|  | di bawah 23 | 17 | 6 | 7 | 4 | 1.5485 |  |
|  | di atas 23 | 5 | 0 | 0 | 5 | 0 |  |
| arah angin |  |  |  |  |  |  | 0 |
|  | barat daya | 22 | 5 | 7 | 9 | 1.5389 |  |
|  | selain barat daya | 0 | 0 | 0 | 0 | 0 |  |

Setelah informasi selesai dihitung, maka diambil nilai informasi *gain* terbesar dari tiga atribut yang ada yaitu atribut suhu minimum sebesar 0.3423, jadi suhu minimum dapat menjadi *node* ke-2.

di atas 80 di bawah 80

**Berawan**

di atas 23 di bawah 23

**Cerah**

Gambar 3.3 Pohon Keputusan Node 2

Pohon keputusan di atas belum terlihat keputusan yang dominan dari setiap cuaca yang dipilih. Maka mencari kembali nilai *entropy* dan *gain* dari setiap atribut suhu minimum adalah di bawah 23 sebagai berikut :

1. Menghitung nilai *entropy* total suhu minimum (di bawah 23)

Berikut ini adalah tabel penyelesaiannya :

Tabel 3.6 Data Atribut Suhu Minimum (di bawah 23)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **No** | **Kota/Kab** | **Suhu Max** | **Arah angin** | **Cuaca** |
|
| 1 | Pandan | Di bawah 80 | Barat Daya/10 KT | Berawan |
| 2 | Kabanjahe | Di bawah 30 | Barat Daya/10 KT | Hujan |
| 3 | Pematang Raya | Di atas 30 | Barat Daya/10 KT | Hujan |
| 4 | Kisaran | Di atas 30 | Barat Daya/10 KT | Cerah |
| 5 | Rantau Prapat | Di atas 30 | Barat Daya/10 KT | Berawan |

Tabel 3.6 Data Atribut Suhu Minimum (di bawah 23) lanjutan

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **No** | **Kota/Kab** | **Suhu Max** | **Arah angin** | **Cuaca** |
|
| 6 | Sidikalang | Di bawah 30 | Barat Daya/10 KT | Hujan |
| 7 | Balige | Di bawah 30 | Barat Daya/10 KT | Hujan |
| 8 | Panyabungan | Di atas 30 | Barat Daya/10 KT | Berawan |
| 9 | Sibolga | Di atas 30 | Barat Daya/10 KT | Berawan |
| 10 | Tebing Tinggi | Di atas 30 | Barat Daya/10 KT | Cerah |
| 11 | Padang Sidempuan | Di atas 30 | Barat Daya/10 KT | Berawan |
| 12 | Kota pinang | Di atas 30 | Barat Daya/10 KT | Berawan |
| 13 | Aek Kanopan | Di atas 30 | Barat Daya/10 KT | Berawan |
| 14 | Lahomi | Di atas 30 | Barat Daya/10 KT | Cerah |
| 15 | Lotu | Di atas 30 | Barat Daya/10 KT | Cerah |
| 16 | Sibuhan | Di atas 30 | Barat Daya/10 KT | Hujan |
| 17 | Gunung Tua | Di atas 30 | Barat Daya/10 KT | Hujan |

*E(S) =*

=

= 1.5485

1. Menghitung *entropy* untuk setiap nilai fitur terhadap kelas. Untuk *entropy* nilai dalam atribut sebagai berikut :
   1. *Entropy* suhu maksimum

*E(S,di bawah 30)*

*=*

=

= 0

*E(S,di atas 30)*

*=*

=

= 1.4926

* 1. *Entropy* arah angin

*E(S,barat daya)*

*=*

=

= 1.5485

*E(S,selain barat daya)*

=

= 0

3. Menghitung nilai *Information Gain*

Menghitung nilai *gain* masing-masing dari tiap atribut adalah sebagai berikut :

1. Menghitung nilai *gain* atribut suhu maksimum

*G(Semua, suhu max)* = *E*(s)  *)*

=1.5485

= 0.3193

1. Menghitung nilai *gain* atribut arah angin

*G(Semua, arah)* = *E*(s)

= 1.5485

= 0

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Node** |  |  | **Jml Kasus (S)** | **Hujan (S1)** | **Berawan**  **(S2)** | **cerah (S3)** | **Entropy** | **Gain** |
| 3 | Suhu min di bawah 23 |  | 17 | 6 | 7 | 4 | 1.5485 |  |
|  | Suhu max |  |  |  |  |  |  | 0.3193 |
|  | di bawah 30 | 3 | 3 | 0 | 0 | 0 |  |
|  | Di atas 30 | 14 | 3 | 7 | 4 | 1.4926 |  |
| arah angin |  |  |  |  |  |  | 0 |
|  | barat daya | 17 | 6 | 7 | 4 | 1.5485 |  |
|  | selain barat daya | 0 | 0 | 0 | 0 | 0 |  |

Tabel 3.7 Hasil Perhitungan Entropy dan Gain untuk Node 3

Setelah informasi selesai dihitung, maka diambil nilai informasi *gain* terbesar dari dua atribut yang ada yaitu atribut suhu maksimum sebesar 0.3193, jadi suhu maksimum dapat menjadi *node* ke-3.

di atas 80 di bawah 80

**Berawan**

di atas 23 di bawah 23

**Cerah**

di atas 30 di bawah 30

**Hujan**

Gambar 3.4 Pohon Keputusan Node 3

Pohon keputusan di atas belum terlihat keseluruhannya hasilnya sehingga perlu mencari kembali nilai *entropy* dan *gain* dari setiap atribut suhu maximum = di atas 30

1. Menghitung nilai *entropy* total suhu maksimum (di atas 30)

Berikut ini adalah tabel penyelesaiannya :

Tabel 3.8 Data Atribut Suhu Maksimum (di atas 30)

|  |  |  |  |
| --- | --- | --- | --- |
| **No** | **Kota/Kab** | **Arah angin** | **Cuaca** |
|
| 1 | Pandan | Barat Daya/10 KT | Berawan |
| 2 | Pematang Raya | Barat Daya/10 KT | Hujan |

Tabel 3.8 Data Atribut Suhu Maksimum (di atas 30)

|  |  |  |  |
| --- | --- | --- | --- |
| **No** | **Kota/Kab** | **Arah angin** | **Keterangan** |
|
| 3 | Kisaran | Barat Daya/10 KT | Cerah |
| 4 | Rantau Prapat | Barat Daya/10 KT | Berawan |
| 5 | Panyabungan | Barat Daya/10 KT | Berawan |
| 6 | Sibolga | Barat Daya/10 KT | Berawan |
| 7 | Tebing Tinggi | Barat Daya/10 KT | Cerah |
| 8 | Padang Sidempuan | Barat Daya/10 KT | Berawan |
| 9 | Kota pinang | Barat Daya/10 KT | Berawan |
| 10 | Aek Kanopan | Barat Daya/10 KT | Berawan |
| 11 | Lahomi | Barat Daya/10 KT | Cerah |
| 12 | Lotu | Barat Daya/10 KT | Cerah |
| 13 | Sibuhan | Barat Daya/10 KT | Hujan |
| 14 | Gunung Tua | Barat Daya/10 KT | Hujan |

*E(S)=*

=

= 1.4926

1. Menghitung *entropy* untuk setiap nilai fitur terhadap kelas. Untuk *entropy* nilai dalam atribut sebagai berikut :
   1. *Entropy* arah angin

*E(S,barat daya) =*

=

= 1.4926

*E(S,selain barat daya)*

=

= 0

3. Menghitung nilai *Information Gain*

Menghitung nilai *gain* masing-masing dari tiap atribut adalah sebagai berikut :

1. Menghitung nilai *gain* atribut arah angin

*G(Semua, arah)* = *E*(s)

= 1.4926

= 0

Tabel 3.9 Hasil Perhitungan Entropy dan Gain untuk Node 4

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Node** |  |  | **Jml Kasus (S)** | **Hujan (S1)** | **Berawan(S2)** | **cerah (S3)** | **Entropy** | **Gain** |
| 4 | suhu max di atas 30 |  | 14 | 3 | 7 | 4 | 1.4926 |  |
|  | arah angin |  |  |  |  |  |  | 0 |
|  |  | barat daya | 14 | 3 | 7 | 4 | 1.4926 |  |
|  |  | selain barat daya | 0 | 0 | 0 | 0 | 0 |  |

Tabel di atas menjelaskan bahwasanya tidak ada lagi *node* yang di proses. Maka *node* keputusan hasil akhir adalah sebagai berikut :

di atas 80 di bawah 80

**Berawan**

di atas 23 di bawah 23

**Cerah**

di atas 30 di bawah 30

**Hujan**

barat daya barat daya

**Berawan**

**Berawan**

Gambar 3.5 Pohon Keputusan Hasil Akhir

* + 1. **Mengubah Tree Menjadi Rule**

Dengan memperhatikan pohon keputusan pada gambar 3.5 diketahui bahwa pohon keputusan telah terbentuk. Berikut ini adalah bentuk *tree* yang diubah menjadi *rule* :

R1 : *If* kelembaban > 80 *Then* cuaca = berawan

R2 : *If* kelembaban <= 80 and suhu minimum > 23 *Then* cuaca = cerah

R3 *: If* kelembaban <= 80 and suhu minimum <= 23 and suhu maksimum <= 30 *Then* cuaca = hujan

R4 *: If* kelembaban <= 80 and suhu minimum <= 23 and suhu maksimum > 30 and arah angin = barat daya *Then* cuaca = berawan

R5 *: If* kelembaban <= 80 and suhu minimum <= 23 and suhu maksimum > 30 and arah angin = barat daya *Then* cuaca = berawan

* 1. **Flowchart Program**

*Flowchart* program merupakan langkah-langkah program yang menceritakan kejadian suatu proses satu dengan proses lainya dalam suatu program secara mendetail yang diwakilkan dalam bentuk simbol atau bagan.

Dalam penelitian ini, akan digunakan program berbasis web dengan *script* *PHP* dan *databse MySQL*. Data yang diperoleh kemudian dimasukkan ke dalam program, setelah data dimasukkan maka data tersebut akan diproses. Seperti yang disajikan pada gambar 3.6 berikut :

Mulai

Masukkan Data Training

Hitung Entropy dan Information Gain dari tiap atribut

Buat simpul akar pohon berdasarkan information gain terbesar

Hitung entropy dan information gain dari tiap atribut dengan menghilangkan atribut yang telah dipilih sebelumya

Buat simpul internal pohon berdasarkan information gain terbesar

Semua atribut sudah masuk pada pohon ?

Selesai

Menghasilkan aturan keputusan

Tidak

Ya

Gambar 3.6 *Flowchart* Program *clasification* Algoritma C4.5

* 1. **Perancangan Sistem**

Perancangan sistem bertujuan untuk membuat suatu pemodelan kerangka dasar *clasification* algoritma C4.5 yang akan digunakan, sistem masukan yang dibutuhkan, keluaran yang diharapkan, serta prosedur penggunaan sistem. Tahapan yang akan dilakukan dalam pemodelan sistem antara lain perancangan *use case diagram, activity diagram, class diagram* dan perancangan *interface.*

**3.4.1 Perancangan *Use Case Diagram***

Perancangan *use case diagram* pada proses penerapan *data mining* dalam memprediksi prakiraan cuaca menggunakan metode Algoritma C4.5adalah sebagai berikut :

Admin

Gambar 3.7 *Use Case* *Diagram* *Clasification* Algoritma C4.5

Tabel 3.4 Skenario *Use Case*

|  |  |
| --- | --- |
| Admin | Sistem |
| Menggunakan sistem Aplikasi*.* Dan memasukkan data ke dalam sistem. |  |
| Menetapkan metode. |  |
|  | Menampilkan proses Algoritma C4.5 |
| Menjalankan/ mengetes data yang sudah dimasukkan. |  |
|  | Memproses perintah *user*. Dan menampilkan hasil data yaitu menampilkan *decision tree* dan *rule* |

**3.4.2 Perancangan *Activity Diagram***

*Activity Diagram*menggambarkan berbagai alir aktivitas dalam sistem yang dipakai, bagaimana masing-masing alir berawal, proses data yang akan terjadi dan bagaimana data tersebut berakhir. *Activity Diagram* juga dapat menggambarkan proses paralel yang mungkin terjadi pada beberapa eksekusi. *Activity Diagaram* merupakan *state* diagram khusus, dimana sebagian besar *state*  adalah *action* dan sebagian besar transisi ditinjau oleh selesainya *s*ebelumnya. Dari *activity diagram* berikut terlihat bagaimana dari proses *decision tree* algoritma C4.5 dalam memprediksi prakiraan cuaca.

|  |  |
| --- | --- |
| *Admin* | Sistem |
| **Mulai**  Pilih menu C4.5  Liat menu C4.5  Menampilkan hasil *decision tree* dan *rule* C4.5  Menampilkan proses Algoritma C4.5  Pilih menu pohon keputusan C4.5  Pilih menu Algoritma C4.5 | Menampilkan menu C4.5  **Selesai** |

Gambar 3.8 *Activity Diagram* Proses *Clasification* Algoritma C4.5

**3.4.3 Perancangan *Class Diagram***

*Class Diagram* merupakan suatu diagram yang menggambarkan atau memvisualisasikan struktur sistem dari kelas-kelas serta hubungannya. Dari *class diagram* berikut terlihat bagaimana dari proses *clasification* algoritma C4.5 dalam memprediksi prakiraan cuaca.

Tbl\_suhu\_min\_≤23

* No\_id :int
* Nama\_kota/kab :varchar
* Kelembaban : varchar
* Arah angin : varchar
* Cuaca : varchar

+ update() : void

+ delete() : void

+ display() : void

Tbl\_normalisasi\_cuaca

* No\_id :int
* Nama\_kota/kab :varchar
* Suhu\_max : varchar
* Suhu\_min : varchar
* Kelembaban : varchar
* Arah\_angin : varchar
* Cuaca : varchar

+ update() : void

+ delete() : void

+ display() : void

Tbl\_suhu\_max\_>30

* No\_id :int
* Nama\_kota/kab :varchar
* Suhu\_min : varchar
* Kelembaban : varchar
* Arah\_angin : varchar
* Cuaca : varchar

+ update() : void

+ delete() : void

+ display() : void

Tbl\_prakiran\_cuaca

* No\_id :int
* Nama\_kota/kab :varchar
* Suh\_ max : int
* Suhu\_min : int
* Kelembaban : int
* Arah\_angin : varchar
* Cuaca : varchar

+ update() : void

+ delete() : void

+ display() : void

Tbl\_kelembaban\_>80

* No\_id :int
* Nama\_kota/kab:varchar
* Arah angin : varchar
* Cuaca : varchar

+ update() : void

+ delete() : void

+ display() : void

Tbl\_arah\_angin

* No\_id :int
* Nama\_kota/kab:varchar
* Cuaca : varchar

+ update() : void

+ delete() : void

+ display() : void

Tbl\_kelembaban\_≤80

* No\_id :int
* Nama\_kota/kab:varchar
* Arah\_angin : varchar
* Cuaca : varchar

+ update() : void

+ delete() : void

+ display() : void

Tbl\_rule

* Suhu\_max:varchar
* Suhu\_min:varchar
* Kelembaban:vrchr
* Arah\_angin:vrchr
* Cuaca:vrchr

+ display() :void

login

* Username : int
* Password : int

+validasi () :void

+ logout() :void

Gambar 3.9 *Class Diagram**Clasification* Algoritma C4.5

**3.4.4 Perancangan** *I****nterface***

Pada perancangan *interface* akan dijelaskan tentang desain *form* dan segmen program. Desain program dijelaskan melalui gambar dan segmen program adalah potongan program dengan *script* PHP*.*

1. Halaman Login

Halaman Login adalah proses untuk mengakses dengan memasukkan identitas dari akun pengguna dan kata sandi. Seperti Gambar 3.10 dengan memasukkan username dan password, setelah itu klik tombol sign in untuk masuk ke halaman home.

**Admin** **Login Form**

Login Admin

Masukkan username dan password

Username............

Password.............

Sign in!

Gambar 3.10 Halaman Login

1. Form Tampilan Home

Rancangan halaman utama dapat dilihat seperti gambar di bawah ini :

**HEADER/JUDUL**

Selamat Datang

Login Hanya untuk Admin

**Copyright ©2016 by xxxx**

* **Home**
* Data Cuaca
* Partisi Data
* C4.5 Proses Mining
* C4.5  Perhitungan
* C4.5  Pohon Keputusan
* Kinerja Tabel Penilaian
* Penentu Pohon Keputusan
* About
* Logout

Gambar 3.11 Form Tampilan Home

Keterangan :

Home : Untuk menampilkan informasi awal

Data Cuaca : Melakukan input data cuaca

Partisi Data : Menampilkan data *training* yang diperoleh

Perhitungan C4.5 : Melakukan proses perhitungan C4.5

* Input data *training*
* Hitung masing-masing nilai atribut
* Hitung nilai *entropy* dari masing-masing nilai
* Hitung nilai *gain* dengan menggunakan nilai *entropy* sebelumya
* Ambil nilai *gain* terbesar untuk jadikan node
* Ulangi perhitungan hingga pohon memiliki kelas

Pohon keputusan C4.5 : Menampilkan pohon keputusan yang dibentuk pohon keputusan berisi *rule-rule*

Tabel penilaian : Menampilkan nilai atribut keputusan dalam proses data min

Penentu pohon keputusan : Untuk melakukan proses penentu keputusan data sehingga dihasilkan hasil keputusan C4.5

About : Informasi pemilik situs

Logout : proses keluar dari halaman website

1. Form Tambah Data Cuaca

**HEADER/JUDUL**

****

**Input Data Cuaca**

**Copyright ©2016 by xxxx**

Kota/Kab:

Suhu max :

Suhu Min :

Kelembaban :

Arah angin :

Cuaca :

* HUJAN

Status Data :

* D.Training

Input

Reset

Gambar 3.12 Form Tampilan Data Cuaca

Gambar 3.12 Form Tampilan Data Cuaca

1. Form Partisi Data

**HEADER/JUDUL**

****

**Data Training - Partisi Data**

**Copyright ©2016 by xxxx**

Proses

Set data training

xx

Status Data Data training Data testing

Cerah xxx xxx

Hujan xxx xxx

Berawan xxx xxx

Gambar 3.13 Form Tampilan Partisi Data

1. Form Proses Mining C4.5

**HEADER/JUDUL**

****

**Proses mining C4.5**

**Copyright ©2016 by xxxx**

Penentu Pohon Keputusan

Pohon keputusan C4.5

Perhitungan C4.5

Proses

Proses

Proses

Gambar 3.12 Form Tampilan Partisi Data

Gambar 3.14 Form Tampilan Proses C4.5