

BAB II

KAJIAN PUSTAKA

Klasifikasi merupakan salah satu metode dalam *data mining* untuk memprediksi label kelas dari suatu *record*. Penelitian yang dilakukan oleh Romansyah tahun 2009 dalam klasifikasi terhadap data diabetes yang suatu pendekatan yang sangat populer dan praktis dalam *machine learning* untuk menyelesaikan permasalahan klasifikasi. Metode ini digunakan untuk memperkirakan nilai diskrit dari fungsi target, yang mana fungsi pembelajaran direpresentasikan oleh sebuah *decision tree*. Sebuah atribut *numeric-symbolic* adalah sebuah atribut yang dapat bernilai *numeric* ataupun *symbolic* yang dihubungkan dengan sebuah variabel kuantitatif. Nilai-nilai seperti ini yang menyebabkan perluasan dari *decision tree*.

Penggunaan teknik *fuzzy* memungkinkan melakukan prediksi suatu objek yang dimiliki oleh lebih dari satu kelas. *Fuzzy decision tree* memungkinkan untuk menggunakan nilai-nilai *numeric-symbolic* selama konstruksi atau saat mengklasifikasikan kasus-kasus baru. Manfaat dari teori himpunan *fuzzy* dalam *decision tree* ialah meningkatkan kemampuan dalam memahami *decision tree* ketika menggunakan atribut-atribut kuantitatif. Bahkan, dengan menggunakan *fuzzy* dapat meningkatkan ketahanan saat melakukan klasifikasi kasus-kasus baru. Dengan menerapkan teknik *data mining* dan *fuzzy* pada data diabetes untuk mendapatkan aturan klasifikasi yang dapat digunakan untuk memprediksi potensi seseorang terserang penyakit diabetes [RSN-09].

2.1. Tebu dan Gula

Tebu adalah tanaman yang ditanam untuk bahan baku gula. Tanaman ini hanya dapat tumbuh di daerah beriklim tropis. Tanaman ini termasuk jenis rumput-rumputan. Umur tanaman sejak ditanam sampai bisa dipanen mencapai kurang lebih 1 tahun. Di Indonesia tebu banyak dibudidayakan di pulau Jawa dan Sumatra. Untuk pembuatan gula, batang tebu yang sudah dipanen diperas dengan mesin pemeras (mesin press) di pabrik gula. Sesudah itu, nira atau air perasan tebu tersebut disaring, dimasak, dan diputihkan sehingga menjadi gula pasir yang

kita kenal. Dari proses pembuatan tebu tersebut akan dihasilkan gula 5%, ampas tebu 90% dan sisanya berupa tetes (molasse) dan air [SOE-11]. Setelah tebu dipanen dan diangkut ke pabrik, selanjutnya dilakukan pengolahan. Pengolahan tebu menjadi gula putih dilakukan di pabrik dengan menggunakan peralatan yang menggunakan peralatan yang sebagian besar bekerja secara otomatis. Beberapa tahap pengolahan, yaitu ekstraksi nira, penjernihan, penguapan, kristalisasi, pemisahan kristal, dan pengeringan, pengemasan serta penyimpanan [TIM-00].

Gula merupakan salah satu komoditas strategis dalam perekonomian Indonesia. Dengan luas areal sekitar 400 ribu ha pada periode 2007-2009, industri gula berbasis tebu merupakan salah satu sumber pendapatan bagi sekitar 900 ribu petani dengan jumlah tenaga kerja yang terlibat mencapai sekitar 1.3 juta orang. Gula juga merupakan salah satu kebutuhan pokok masyarakat, maka dinamika harga gula akan mempunyai pengaruh langsung terhadap laju inflasi. [SOE-11].

Hal penting yang perlu diperhatikan dalam pengembangan usaha industri gula berbasis tebu antara lain adalah [SOE-11]:

- Pengelolaan pada aspek on-farm yakni penerapan kaidah teknologi pertanian yang baik dan benar mulai dari persiapan lahan, pengolahan dan penanaman yang mengikuti kaidah masa tanam optimal.
- Pemilihan dan komposisi varietas bibit unggul bermutu.
- Penggunaan, pemeliharaan serta tebang angkut muat (panen).

Gula berasal dari tebu yang memiliki banyak varietas. Berbagai varietas tersebut dapat dibedakan melalui beberapa faktor antara lain, batang tebu, jumlah ruas batang tebu dan kandungan cairan yang terdapat dalam tebu, seperti % brix, Harkat Kemurnian (HK), nira dan rendemen.

2.2. Faktor Pembeda Varietas Tebu

2.2.1. Batang

Pada pertumbuhan batang yang baik akan didapat tinggi batang 3-5 meter atau lebih. Bagian luar merupakan kulit keras, sedang bagian dalam lunak dan mengandung nira (air gula) [BAS-10]. Untuk kepentingan pengolahan gula, batang tanaman tebu dianggap tersusun atas nira tebu dan ampas.

2.2.2. Ruas Batang Tebu

Perbedaan varietas tebu menyebabkan perbedaan tebal batang, warna, panjang ruas, bentuk ruas, bentuk mata dan lapisan lilin. Kedudukan ruas satu dengan lain: tegak atau zig zag, sedang bentuk ruas dapat bervariasi sesuai varietasnya. Panjang ruas makin panjang dari pangkal sampai pertengahan batang tebu dan selanjutnya makin pendek ke ujung tanaman tebu. Jika tanaman tebu mengalami kekurangan air dan kemudian pada musim hujan tiba mendapat cukup air, maka terdapat ruas-ruas yang pendek dan kemudian di atasnya ruas-ruas yang panjang. Hal ini dapat menyebabkan tebu roboh [BAS-10].

Mata ruas tebu adalah kuncup tebu yang terletak pada buku ruas batang dan terlindung oleh pangkal pelepah daun. Bagian ini merupakan salah satu tanda pengenal varietas [BAS-10].

2.2.3. Persentase (%) Brix

Brix merupakan urutan kepekaan larutan gula yang didefinisikan sebagai jumlah (gram) zat padat total terlarut per 100 gram larutan gula. Contoh %Brix: Pembacaan Brix 15,5 suhu terbaca 27°C Brix terkoreksi $(15,5+0)=15,5$ [BAS-10]. Brix diukur dengan menggunakan refraktometer [SUP-07].

2.2.4. HK

HK (Harkat Kemurnian) merupakan perbandingan antara pol dan brix. Pol didefinisikan sebagai jumlah (gram) gula yang terlarut setiap 100 gram larutan gula [BAS-10].

2.2.5. Nira

Nira merupakan cairan yang terdapat dalam batang tebu [BAS-10]. Selain sukrosa, nira tebu juga mengandung bahan pengotor berupa bahan-bahan non-gula hingga 50%, seperti gula pereduksi, asam organik, asam an- organik, asam amino, protein, pati, lilin, gum, mineral (seperti kalium, magnesium, kalsium, dan silika), bahan pembentuk warna (klorofil), dan bahan ter-suspensi lainnya [SUP-07].

Pada proses penggilingan tebu, terdapat lima kali proses penggilingan dari batang tebu sampai dihasilkan ampas tebu. Pada penggilingan pertama dan kedua

dihasilkan nira mentah yang berwarna kuning kecoklatan, kemudian pada proses penggilingan ketiga, keempat dan kelima dihasilkan nira dengan volume yang tidak sama. Setelah proses penggilingan awal yaitu penggilingan pertama dan kedua dihasilkan ampas tebu basah. Untuk mendapatkan nira yang optimal, pada penggilingan ampas hasil gilingan kedua harus ditambahkan susu kapur 3Be yang berfungsi sebagai senyawa yang mampu menyerap nira dari serat ampas tebu, sehingga pada penggilingan ketiga nira masih dapat diserap meskipun volumenya lebih sedikit dari hasil gilingan kedua. Pada penggilingan seterusnya hingga penggilingan kelima ditambahkan susu kapur 3Be dengan volume yang berbeda-beda tergantung sedikit banyaknya nira yang masih dapat dihasilkan [SIR-10].

2.2.6. Rendemen

Rendemen adalah gula yang dihasilkan dari proses pengolahan nira tebu di suatu Pabrik Gula. Faktor-faktor yang mempengaruhi rendemen adalah On Farm (kualitas tebu, varietas, kadar nira tebu, manajemen tebang angkut) dan Off Farm (efisiensi pabrik) [BAS-10].

Rendemen didekati dengan rumus [EMB-09]:

- a) Rendemen = Nilai nira perahan pertama (NNPP) x Kadar nira tebu (KNT) x Efisiensi pabrik
- b) Rendemen = pol tebu x efisiensi pabrik

Kedua rumus ini menghasilkan angka rendemen yang sama. Rumus ini sangat tepat untuk penilaian rendemen nyata dari tebu yang telah berada di emplasemen pabrik, karena sistem bagi hasil rendemen didasarkan pada tebu yang sedang diolah tersebut. Sehingga apabila target produksi (khususnya rendemen) yang telah dicanangkan untuk pabrik gula tersebut tidak tercapai dikatakan bahwa penyebabnya adalah efisiensi pabrik yang rendah [EMB-09].

Usahakan agar tebu ditebang saat rendemen pada posisi optimal yaitu sekitar bulan Agustus atau tergantung jenis tebu. Tebu yang berumur 10 bulan akan mengandung saccharose 10 %, sedang yang berumur 12 bulan bisa mencapai 13 % [SOE-11].

2.3. Varietas Tebu

Ada berbagai macam varietas unggul tebu yang dapat dibudidayakan petani, antara lain PS851, PS862, PS863, PS864, PSBM901, PS921, Bululawang, PSCO902, PSJT941, Kidang Kencana, PS865, PS881, PS882 dan varietas Kentung yang merupakan varietas-varietas unggulan dengan kategori pengelompokan masak awal, masak tengah dan masak akhir sebagai salah satu penerapan manajemen pembibitan untuk menyelaraskan pelaksanaan tertib tanam dan panen [SOE-11].

Jenis bibit sangat menentukan keberhasilan budidaya Tanaman Tebu, karena dengan menggunakan jenis bibit yang tepat akan menghasilkan produktivitas sesuai dengan potensi bibit. Berdasar tingkat kemasakan jenis bibit dapat digolongkan menjadi [BAS-10]:

- a. Jenis Bibit Masak Awal, adalah bibit yang mempunyai sifat genetik sangat peka terhadap sifat kekeringan tanah (lengas tanah) pada kondisi ketersediaan air dalam tanah turun dari $\pm 50\%$ menjadi $\pm 40\%$ kapasitas lapang. Saat itu terjadi pengisian sukrosa. Kondisi ini ditandai dengan tebu berbunga. Varietas tebu masak awal adalah PS862, PS 851, PS 881, persentase masak awal disesuaikan kondisi tipologi wilayah.
- b. Jenis Bibit Masak Tengah, adalah bibit yang mempunyai sifat genetik kepekaan terhadap sifat kekeringan tanah (lengas tanah) kondisi ketersediaan air dalam tanah turun dari $\pm 50\%$ menjadi $\pm 30\%$ kapasitas lapang. Saat itu terjadi pengisian sukrosa. Kondisi ini ditandai dengan tebu berbunga secara sporadis. Varietas tebu masak tengah adalah PSJT 941, KK, PS 921, PS 864, HW, persentase masak tengah sesuai tipologi wilayah.
- c. Jenis bibit masak lambat, adalah bibit yang mempunyai sifat genetik kepekaan terhadap sifat kekeringan tanah (lengas tanah) kondisi ketersediaan air dalam tanah turun dari $\pm 50\%$ menjadi $\pm 25\%$ kapasitas lapang. Saat itu terjadi pengisian sukrosa. Varietas tebu masak lambat adalah PS 864, BL, PS 951 persentase masak awal disesuaikan kondisi tipologi wilayah.

Pada skripsi ini akan dilakukan penelitian klasifikasi pada beberapa varietas tebu, yaitu PS 862, PS 881, Bululawang dan HW. Untuk melakukan klasifikasi tersebut salah satu metodenya adalah menggunakan metode *Fuzzy Iterative Dichotomiser 3* (FID 3). Berikut akan dijelaskan metode klasifikasi yang digunakan, sebelum membahas *Iterative Dichotomiser* (ID3) akan dijelaskan logika fuzzy terlebih dahulu.

2.4. Logika Fuzzy

Logika fuzzy merupakan salah satu komponen pembentuk *softcomputing*. Logika fuzzy pertama kali diperkenalkan oleh Prof. Lotfi A. Zadeh pada tahun 1965. Dasar logika fuzzy adalah teori himpunan fuzzy. Pada teori himpunan fuzzy, peranan derajat keanggotaan sebagai penentu keberadaan elemen dalam suatu himpunan sangatlah penting. Nilai keanggotaan atau derajat keanggotaan atau *membership function* menjadi ciri utama dari penalaran dengan logika fuzzy tersebut [KUS-10].

Pada himpunan tegas (*crisp*), nilai keanggotaan suatu item x dalam suatu himpunan A , yang sering ditulis dengan $\mu_A(x)$, memiliki dua kemungkinan, yaitu [KUS-10]:

- Satu (1), yang berarti bahwa suatu item menjadi anggota dalam suatu himpunan, atau
- Nol (0), yang berarti bahwa suatu item tidak menjadi anggota dalam suatu himpunan.

Misalkan U adalah *universe* (semesta) objek dan x adalah anggota U . Suatu *fuzzy set* A di dalam U didefinisikan sebagai suatu fungsi keanggotaan $\mu_A(x)$, yang memetakan setiap objek di U menjadi suatu nilai real dalam interval $[0,1]$. Nilai-nilai $\mu_A(x)$ menyatakan derajat keanggotaan x di dalam A [SUY-11]. Himpunan fuzzy memiliki 2 atribut, yaitu [KUS-10]:

- a. Linguistik, yaitu penanaman suatu grup yang mewakili suatu keadaan atau kondisi tertentu dengan menggunakan bahasa alami.
- b. Numeris, yaitu nilai (angka) yang menunjukkan ukuran dari suatu variabel.

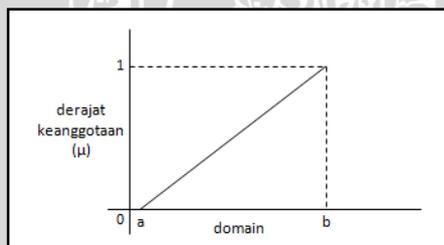
2.4.1. Fungsi Keanggotaan

Fungsi keanggotaan (*membership function*) adalah suatu kurva yang menunjukkan pemetaan titik-titik *input* data ke dalam nilai keanggotaannya (sering disebut dengan derajat keanggotaan) yang memiliki interval antara 0 sampai 1. Salah satu cara yang dapat digunakan untuk mendapatkan nilai keanggotaan adalah melalui pendekatan fungsi. Ada beberapa fungsi yang bisa digunakan, yaitu representasi linear, representasi kurva segitiga, representasi kurva trapesium, representasi kurva bentuk bahu, representasi kurva-S dan representasi kurva bentuk lonceng (*Bell Curve*). Berikut beberapa fungsi keanggotaan yang digunakan [KUS-10]:

a. Representasi Linier

Pada representasi linier, pemetaan input ke derajat keanggotaannya digambarkan sebagai suatu garis lurus. Terdapat 2 keadaan himpunan fuzzy yang linier:

1. Kenaikan himpunan dimulai pada nilai domain yang memiliki derajat keanggotaan nol (0) bergerak ke kanan menuju ke nilai domain yang memiliki derajat keanggotaan lebih tinggi. Gambar 2.1 merupakan representasi linier naik.



Gambar 2.1 Representasi Linear Naik

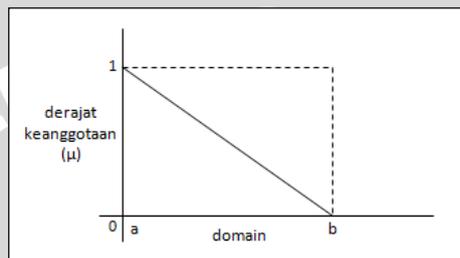
Sumber : [KUS-10]

Kurva tersebut merupakan representasi Linear Naik dengan a merupakan nilai domain dengan derajat keanggotaan 0 ($\mu = 0$) dan bergerak ke kanan menuju b . Dimana b merupakan nilai domain tertinggi dengan derajat keanggotaan 1 ($\mu = 1$). Untuk mencari derajat keanggotaan suatu titik mulai dari a sampai dengan b dapat dirumuskan menggunakan fungsi keanggotaan berikut:

Fungsi keanggotaan:

$$\mu[x] = \begin{cases} 0; & x \leq a \\ \frac{(x-a)}{(b-a)}; & a \leq x \leq b \\ 1; & x \geq b \end{cases} \quad (2-1)$$

2. Garis lurus dimulai dari nilai domain dengan derajat keanggotaan tertinggi pada sisi kiri, kemudian bergerak menurun ke nilai domain yang memiliki derajat keanggotaan lebih rendah. Gambar 2.2 merupakan representasi linier turun.



Gambar 2.2 Representasi Linear Turun

Sumber : [KUS-10]

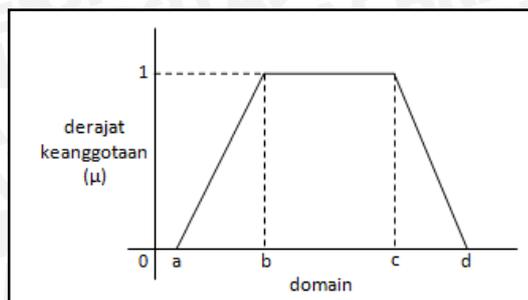
Kurva tersebut merupakan representasi Linear Turun dengan a merupakan nilai domain dengan derajat keanggotaan 1 ($\mu = 1$) dan bergerak ke kanan menuju b . Dimana b merupakan nilai domain terendah dengan derajat keanggotaan 0 ($\mu = 0$). Untuk mencari derajat keanggotaan suatu titik mulai dari a sampai dengan b dapat dirumuskan menggunakan fungsi keanggotaan berikut.

Fungsi keanggotaan:

$$\mu[x] = \begin{cases} \frac{b-x}{b-a}; & a \leq x \leq b \\ 0; & x \geq b \end{cases} \quad (2-2)$$

- b. Representasi Kurva Trapesium

Kurva trapesium pada dasarnya seperti bentuk segitiga, hanya saja terdapat beberapa titik yang memiliki nilai keanggotaan 1. Gambar 2.3 merupakan kurva trapesium.



Gambar 2.3 Kurva Trapesium

Sumber : [KUS-10]

Pada kurva trapesium terdapat 4 titik, dimana titik a dan d merupakan nilai domain terendah dengan derajat keanggotaan 1 ($\mu=0$), sedangkan titik b dan c merupakan nilai domain tertinggi dengan derajat keanggotaan 1 ($\mu=1$). Dimana titik a bergerak menuju titik b, yang mana merupakan representasi linier naik. Titik c beregerak menuju titik d, yang mana merupakan representasi linier turun. Untuk mencari derajat keanggotaan dari titik a hingga titik d dapat dirumuskan menggunakan fungsi keanggotaan berikut.

Fungsi Keanggotaan:

$$\mu[x] = \begin{cases} 0; & x \leq a \text{ atau } x \geq d \\ \frac{(x-a)}{(b-a)}; & a \leq x \leq b \\ 1; & b \leq x \leq c \\ \frac{d-x}{d-c}; & c \leq x \leq d \end{cases} \quad (2-3)$$

Fuzzy Inference System Mamdani merupakan metode yang digunakan untuk menguji data tebu. Pengujian ini dilakukan agar mengetahui keakuratan sistem. Metode tersebut akan dibahas sebagai berikut:

2.4.2. *Fuzzy Inference System Mamdani*

Sistem inferensi *fuzzy* adalah suatu *framework* yang didasarkan pada konsep himpunan *fuzzy*, *fuzzy if-then rules*, dan *fuzzy reasoning*. Salah satu metode inferensi *fuzzy* yang paling umum digunakan adalah metode sistem inferensi *fuzzy* Mamdani. Struktur dasar dari sistem inferensi *fuzzy* terdiri dari tiga komponen yaitu [JSM-97]:

1. *Rule base*, terdiri dari aturan-aturan *fuzzy* (*fuzzy rules*)
2. *Database/dictionary*, mendefinisikan fungsi keanggotaan yang digunakan pada aturan *fuzzy*, dan
3. *Reasoning mechanism*, melakukan proses inferensi pada aturan dan fakta yang diberikan untuk memperoleh output atau kesimpulan.

Metode *Mamdani* sering dikenal sebagai metode *Max-Min*. metode ini diperkenalkan oleh Ebrahim Mamdani pada tahun 1975. Untuk mendapatkan output diperlukan 4 tahapan [KUS-10]:

1. Pembentukan himpunan *fuzzy*

Pada Metode *Mamdani*, baik variabel *input* maupun variabel *output* dibagi menjadi satu atau lebih himpunan *fuzzy*.

2. Aplikasi fungsi implikasi

Pada Metode *Mamdani*, fungsi implikasi yang digunakan adalah *Min*. Fungsi ini akan memotong *output* himpunan *fuzzy*.

3. Komposisi aturan

Tidak seperti penalaran monoton, apabila sistem terdiri dari beberapa aturan, maka inferensi diperoleh dari kumpulan dan korelasi antar-aturan. Ada 3 metode yang digunakan dalam melakukan inferensi sistem *fuzzy*, yaitu: *max*, *additive* dan *probabilistic OR* (*probor*).

Pada metode *max*, solusi himpunan *fuzzy* diperoleh dengan cara mengambil nilai maksimum aturan, kemudian menggunakannya untuk memodifikasi daerah *fuzzy* dan mengaplikasikannya ke *output* dengan menggunakan operator *OR*. Jika semua proposisi telah diaplikasikan maka *output* akan berisi suatu himpunan *fuzzy* yang merefleksikan kontribusi dari tiap-tiap proposisi. Secara umum dapat dituliskan :

$$\mu_{sf}[x_i] = \text{Max}(\mu_{sf}[x_i], \mu_{kf}[x_i]) \quad (2-4)$$

Dengan :

$\mu_{sf}[x_i]$ = nilai keanggotaan solusi *fuzzy* sampai aturan ke-*i*

$\mu_{kf}[x_i]$ = nilai keanggotaan konsekuen *fuzzy* aturan ke-*i*.

4. Penegasan (*defuzzy*)

Input dari proses *defuzzifikasi* adalah suatu himpunan *fuzzy* yang diperoleh dari komposisi aturan-aturan *fuzzy*, sedangkan *output* yang dihasilkan

merupakan suatu bilangan pada domain himpunan *fuzzy* tersebut. Untuk *hard classification* berlaku operasi MAX pada keluaran defuzzifikasi. Pola tersebut disusun untuk *class c* dengan nilai *membership class* yang tertinggi. Secara matematika dirumuskan:

$$\forall j \in 1, 2, \dots, c \text{ and } j \neq c, \quad F_c(x) \geq F_j(x), \quad (2-6)$$

Dimana $F_j(x)$ merupakan nilai *membership* untuk *class j*.

Operasi defuzzifikasi MAX digunakan untuk masalah klasifikasi agar mendapatkan label *hard class* sebagai *class* klasifikasinya [GMS-07].

Fuzzy Iterative Dichotomiser 3 merupakan salah satu metode klasifikasi gabungan 2 metode, yaitu logika fuzzy dan *Iterative Dichotomiser 3* (ID3). Atribut-atribut pada klasifikasi tersebut akan diproses terlebih dahulu menggunakan himpunan *fuzzy* kemudian akan diproses dalam ID3. Dimana ID3 merupakan salah satu metode klasifikasi dalam data mining. Dan ID3 merupakan salah satu metode yang menggunakan teknik *decision tree*. Metode-metode tersebut akan dibahas sebagai berikut, dan yang akan dibahas terlebih dahulu adalah data mining.

2.5. Data Mining

Data mining merupakan proses untuk menemukan pengetahuan (*knowledge discovery*) yang ditambang dari sekumpulan data yang volumenya sangat besar. Data mining mengacu pada proses untuk menambang (*mining*) pengetahuan dari sekumpulan data yang sangat besar. Sebenarnya data mining merupakan suatu langkah dalam *knowledge discovery in databases* (KDD) [AYU-07]. Model data mining dibuat berdasarkan salah satu dari dua jenis pembelajaran, yaitu *supervised* dan *unsupervised*. Fungsi pembelajaran *supervised* digunakan untuk memprediksi suatu nilai. Fungsi pembelajaran *unsupervised* digunakan untuk mencari struktur intrinsik, relasi dalam suatu data yang tidak memerlukan class atau label sebelum dilakukan proses pembelajaran. Contoh dari algoritma pembelajaran *supervised* yaitu *Decision tree* untuk klasifikasi.

Klasifikasi (*Classification*) 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

[HAK-01]. Teknik klasifikasi bekerja dengan mengelompokkan data berdasarkan data *training* dan nilai atribut klasifikasi. Aturan pengelompokan tersebut akan digunakan untuk klasifikasi data baru ke dalam kelompok yang ada. Klasifikasi dapat direpresentasikan dalam bentuk pohon keputusan (*decision tree*). Setiap node dalam pohon keputusan menyatakan suatu tes terhadap atribut dataset, sedangkan setiap cabang menyatakan hasil dari tes tersebut. Pohon keputusan yang terbentuk dapat diterjemahkan menjadi sekumpulan aturan dalam bentuk IF *condition* THEN *outcome*. [AYU-07].

2.5.1. Decision Tree

Decision tree merupakan salah satu fungsional dari data mining yang menggunakan representasi tree untuk menentukan aturan-aturan klasifikasi. *Decision tree* dapat juga dikatakan sebagai flowchart seperti struktur tree, dimana tiap node internal menunjukkan sebuah tes pada sebuah atribut, tiap cabang menunjukkan hasil dari tes, dan setiap node leaf menunjukkan kelas-kelas atau distribusi kelas. Metode *Decision tree* sangat terkenal daripada metode klasifikasi yang lainnya, karena metode ini tidak membutuhkan pengetahuan yang lebih atau pengaturan parameter [HAK-01]. *Decision tree* digunakan untuk memperkirakan nilai diskret dari fungsi target, yang mana fungsi pembelajaran direpresentasikan oleh sebuah *decision tree* atau pohon keputusan [LIA-05].

Ada berbagai pohon keputusan *top-down decision tree* seperti ID3 (Quinlan,1986), C4.5 (Quinlan, 1993), CART (Breiman et al., 1984).Beberapa terdiri dari dua fase konseptual: tumbuh dan pemangkasan (C4.5 dan CART). Lainnya hanya melakukan fase tumbuh.Algoritma *greedy* dengan alami dan membangun pohon keputusan pada top down, cara rekursif (juga dikenal dengan *divide* dan *conquer*). Dalam setiap iterasi, algoritma mempertimbangkan partisi dari himpunan *training* menggunakan hasil dari fungsi diskritdari atribut masukan. Pemilihan fungsi yang paling tepat dibuat menurut beberapa langkah pemecahan. Setelah pemilihan yang sesuai dibagi, masing-masing node lanjut membagi *training* ditetapkan menjadi subset yang lebih kecil, sampai memperoleh ukuran pemisahan cukup atau berhenti pada kriteria yang terpenuhi [MAR-10].

2.5.2. Iterative Dichotomiser 3 (ID3)

Iterative Dichotomiser 3 (ID3) merupakan salah satu algoritma yang paling sering digunakan pada mesin *learning* dan data mining karena kemudahan untuk penggunaan dan efektivitas. J. Rose Quinlan mengembangkannya pada tahun 1986 berdasarkan pada algoritma Concept Learning System (CLS). Algoritma ini membangun *decision tree* dari beberapa data dalam proses pengklasifikasian dan memprediksi klasifikasi data baru. Data harus memiliki beberapa atribut dengan nilai yang berbeda. Sementara itu, data juga memiliki beragam standar, kelas diskrit (yaitu ya / tidak). *Decision tree* memilih atribut untuk pengambilan keputusan dengan menggunakan *Information Gain* (IG) [LIA-05].

Berikut algoritma ID3 [SUY-11]:

Function ID3 (*KumpulanSampel*, *AtributTarget*, *KumpulanAtribut*)

1. Buat simpul *Root*.
2. **If** semua sampel adalah kelas i , maka **Return** pohon satu simpul *Root* dengan label = i .
3. **If** *KumpulanAtribut* kosong, **Return** pohon satu simpul *Root* dengan label = nilai atribut target yang paling umum (yang paling sering muncul)

Else

- $A \leftarrow$ Atribut yang merupakan *the best classifier* (dengan *information gain* terbesar)
- Atribut keputusan untuk *Root* $\leftarrow A$
- **For** v_i (setiap nilai pada A)
 - Tambahkan suatu cabang di bawah *Root* sesuai dengan nilai v_i
 - Buat suatu variable, misalnya $Sampel_{v_i}$, sebagai himpunan bagian (subset) dari *KumpulanSampel* yang bernilai v_i pada atribut A
 - **If** $Sampel_{v_i}$ kosong
 - **Then** dibawah cabang ini ditambahkan suatu simpul daun (*leaf node*, simpul yang tidak punya anak dibawahnya) dengan label = nilai atribut target yang paling umum (yang paling sering muncul)
 - **Else** di bawah cabang ini ditambahkan *subtree* dengan memanggil fungsi ID3($Sampel_{v_i}$, *AtributTarget*, *Atribut*-[A])

End

End

End

4. Return Root

Setelah dilakukan perhitungan menggunakan ID3, kemudian akan dilakukan perhitungan menggunakan *fuzzy decision tree*. Dimana didalamnya terdapat *threshold* yang merupakan salah satu metode untuk mengekspansi node pada salah satu atributnya untuk membentuk *tree*. Serta *fuzzy entropy* dan *information gain* yang merupakan perhitungan yang dilakukan untuk mencari atribut mana yang akan diekspansi.

2.6. Fuzzy Decision Tree

Fuzzy decision tree merupakan suatu pendekatan yang sangat populer dan praktis dalam *machine learning* untuk menyelesaikan permasalahan klasifikasi yang mengalami ketidakpastian. *Fuzzy decision tree* memungkinkan untuk menggunakan nilai-nilai *numeric-symbolic* selama konstruksi atau saat mengklasifikasikan kasus-kasus baru. Manfaat dari teori himpunan fuzzy dalam *decision tree* ialah meningkatkan kemampuan atribut-atribut kuantitatif. bahkan dengan menggunakan teknik fuzzy dapat meningkatkan ketahanan saat melakukan klasifikasi kasus-kasus baru. [RSN-09].

Salah satu algoritma yang dapat digunakan untuk membangun fuzzy *decision tree* adalah algoritma *Iterative Dichotomiser 3* (ID3). *Iterative Dichotomiser 3* (ID3) adalah algoritma *decision tree learning* (algoritma pembelajaran pohon keputusan) yang paling dasar. Algoritma ini melakukan pencarian secara rakus /menyeluruh (*greedy*) pada semua kemungkinan pohon keputusan. ID3 dikembangkan oleh J. Ross Quinlan. Algoritma ID3 dapat diimplementasikan menggunakan fungsi *rekursif* (fungsi yang memanggil dirinya sendiri). Algoritma ID3 berusaha membangun *decision tree* [WAH-09].

2.6.1. Threshold dalam Fuzzy Decision Tree

Pada proses konstruksi *decision tree* dilakukan proses *pruning* dengan menggunakan *Threshold* dalam *Fuzzy Decision Tree* (FDT). Jika pada proses

learning dari FDT dihentikan sampai semua data contoh pada masing-masing *leaf-node* menjadi anggota sebuah kelas, akan dihasilkan akurasi yang rendah. Oleh karena itu, untuk meningkatkan akurasinya, proses *learning* harus dihentikan lebih awal atau melakukan pemotongan tree secara umum. Untuk itu diberikan 2 *threshold* yang harus terpenuhi jika tree akan diekspansi, yaitu [LIA-05] :

1. *Fuzziness Control Threshold* (FCT) / θ_r

Jika proporsi dari himpunan data dari kelas C_k lebih besar atau sama dengan nilai *threshold* θ_r , maka hentikan ekspansi tree. Sebagai contoh: jika pada sebuah *sub-dataset* rasio dari kelas 1 adalah 90%, kelas 2 adalah 10% dan θ_r adalah 85% maka hentikan ekspansi *tree*

2. *Leaf Decision Threshold* (LDT) / θ_n

Jika banyaknya anggota himpunan data pada suatu *node* lebih kecil dari *threshold* θ_n , hentikan ekspansi tree. Sebagai contoh, sebuah himpunan data memiliki 600 contoh dengan θ_n adalah 2%. Jika jumlah data contoh pada sebuah *node* lebih kecil dari 12 (2% dari 600), maka hentikan ekspansi *tree*.

2.6.2. *Fuzzy Entropy dan Information Gain*

Untuk menghitung *information gain*, terlebih dahulu harus memahami suatu ukuran lain yang disebut *entropy*. Di dalam bidang *Information Theory*, sering menggunakan *entropy* sebagai suatu parameter untuk mengukur heterogenitas (keberagaman) dari suatu kumpulan sampel data. Jika kumpulan sampel data semakin heterogen, maka nilai *entropy*-nya semakin besar [SUY-11].

Setelah mendapatkan nilai *entropy* untuk suatu kumpulan sampel data, maka kita dapat mengukur efektivitas suatu atribut dalam mengklasifikasikan data. Ukuran efektivitas ini disebut *information gain* [SUY-11].

Pada himpunan data fuzzy terdapat penyesuaian rumus untuk menghitung nilai *entropy* untuk atribut dan *information gain* karena adanya ekspresi data fuzzy. Persamaan berikut adalah persamaan untuk mencari nilai fuzzy *entropy* dari keseluruhan data:

$$H_f(S) = H_s(s) = -\sum_i^N P_i * \log_2(P_i) \quad (2-7)$$

Dimana P_i adalah probabilitas dari kelas C_i pada contoh $S=\{x_1, x_2, \dots, x_k\}$.

Untuk menentukan fuzzy *entropy* dari setiap atribut *value* dan *information gain* dari suatu atribut digunakan Persamaan (2-8) dan (2-9).

$$H_f(S, A) = - \sum_{i=1}^c \frac{\sum_j^N \mu_{ij}}{s} \log_2 \frac{\sum_j^N \mu_{ij}}{s} \quad (2-8)$$

$$G_f(S, A) = H_f(S) - \sum_{v \subseteq A} \frac{|S_v|}{|S|} * H_f(S_v, A) \quad (2-9)$$

$H_f(S, A)$ menunjukkan *entropy* atribut A pada himpunan contoh S, dimana kelas target adalah $C_i (i=1..c)$ dan μ_{ij} adalah nilai keanggotaan dari contoh ke-j untuk kelas ke-i. $G_f(S, A)$ menunjukkan *information gain* atribut A pada himpunan contoh S. $|S_v|$ adalah ukuran dari subset $S_v \subseteq S$ pada data pelatihan x_j dengan atribut *value* v. $|S|$ menunjukkan jumlah data sampel dari himpunan S [RSN-09].

2.6.3. Pembangunan Fuzzy Decision Tree dengan Iterative Dichotomiser 3 (ID3)

Misalkan diberikan data set D dengan atribut A_1, A_2, \dots, A_n dan kelas target $C = (C_1, C_2, \dots, C_n)$. D^{ck} merupakan fuzzy subset dari D dengan kelas C_k , dan $|D|$ merupakan jumlah dari nilai keanggotaan pada fuzzy set pada data D. Algoritma fuzzy ID3 *decision tree* adalah sebagai berikut [ABU-09].

1. Bangkitkan root *node* yang memiliki himpunan data fuzzy dengan nilai keanggotaan 1
2. *Node* adalah *leaf*, jika fuzzy set pada data memenuhi kondisi berikut :
 - Jumlah objek pada data set dalam *node* data set D kurang dari Θ_n yang diberikan, $|D| < \Theta_n$.
 - Proporsi data set pada setiap kelas $C_k |D^{ck}|$ pada *node* data set D lebih besar dari atau sama dengan $\Theta_r, \frac{|D^{ck}|}{|D|} \geq \Theta_r$
 - Tidak ada atribut yang tersedia.
3. *Node leaf* diinisialisasikan sebagai nama kelas dengan nilai keanggotaan terbesar atau *node leaf* diinisialisasikan dengan semua kelas beserta nilai keanggotaannya.
4. Jika *node* tidak memenuhi kondisi di atas, maka ikuti langkah berikut:
 - Untuk semua atribut, hitung fuzzy *entropy* dan *information gain*, pilih atribut yang memiliki nilai *information gain* terbesar sebagai atribut tes.

- Bagilah fuzzy data set pada *node* ke dalam fuzzy subset menggunakan atribut tes yang terpilih, dengan nilai keanggotaan sebuah objek yang ada dalam subset set merupakan nilai keanggotaan yang diproduksi *parent node* data set dan nilai dari atribut fuzzy *term* yang terpilih.
 - Untuk setiap subset, bangkitkan node baru dengan label pada cabang merupakan fuzzy *term* dari atribut yang terpilih
5. Untuk setiap *node* baru yang dibangkitkan ulangi langkah 2 secara rekursif.

2.7. Variabel Fuzzy Klasifikasi Tebu Studi Kasus PG. Semboro

Data Klasifikasi Tebu Studi Kasus PG. Semboro merupakan data analisa kemasakan tebu pada tahun 2012-2013 yang dilakukan oleh karyawan Penelitian dan Pengembangan (Litbang) PG. Semboro. Data ini terdiri dari 8 atribut, yaitu panjang tebu (m), diameter tebu (cm), banyak ruas per tebu, berat per meter (kg), persentase (%) brix, harkat kemurnian (HK), nira dan rendemen. Variabel *fuzzy* ini sebelumnya telah dikonsultasikan dengan pakar tanaman tebu yaitu Bapak Ir. Bambang Hadi S. sebagai Staf Kepala II Bidang Tanaman.

Atribut panjang tebu (m), diameter tebu (cm), banyak ruas per tebu, berat per meter (kg), persentase (%) brix, harkat kemurnian (HK), nira dan rendemen memiliki atribut *value* yang berbeda-beda, kecuali atribut panjang tebu (m) dan diameter tebu (cm) serta dibagi menjadi 3 kelompok yaitu buruk, baik dan sangat baik. Dari pembagian itu dapat ditentukan nilai derajat keanggotaan dari himpunan *fuzzy* buruk, baik, dan sangat baik.

Berdasarkan penilaian pakar, diketahui bahwa batasan nilai yang diberikan pada masing-masing anggota himpunan tidak tumpang tindih satu sama lain. Oleh karena itu, pada pembentukan domain fuzzy dibuat *overlapping* dengan dasar *range* pakar. *Overlapping* merupakan kriteria penting dalam pembuatan fungsi keanggotaan [SIV-07]. Nilai *overlap* tidak lebih dari 20% batasan pakar. Jika *overlap* lebih dari 20% mengakibatkan batasan domain *fuzzy* melebihi batasan dari pakar. Misal, panjang tebu memiliki batasan baik ($2,5 \leq \text{panjang} < 3$), jika *overlap* 25% maka $2,5 + 0,625$ (25% dari 2,5) = 3,125 dan itu merupakan batasan sangat baik.

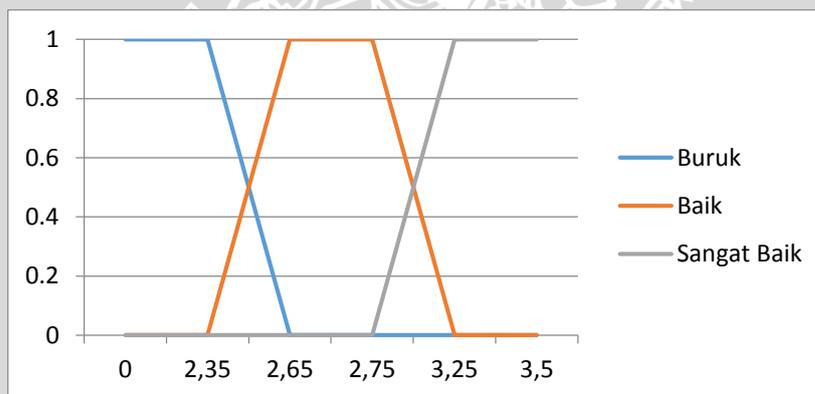
Atribut panjang (m) tebu memiliki *range*, yaitu buruk (panjang < 2,5), baik (2,5 ≤ panjang < 3) dan sangat baik (panjang ≥ 3). Berikut merupakan himpunan *fuzzy* atribut panjang (m) tebu:

$$\mu_{buruk}(x) = \begin{cases} 1 & ; x < 2,35 \\ \frac{2,65 - x}{(2,65 - 2,35)} & ; 2,35 \leq x \leq 2,65 \\ 0 & ; x > 2,65 \end{cases}$$

$$\mu_{baik}(x) = \begin{cases} 0 & ; x < 2,35 \vee x > 3,25 \\ \frac{x - 2,35}{(2,65 - 2,35)} & ; 2,35 \leq x \leq 2,65 \\ 1 & ; 2,65 < x < 2,75 \\ \frac{3,25 - x}{(3,25 - 2,75)} & ; 2,75 \leq x \leq 3,25 \end{cases}$$

$$\mu_{sangat\ baik}(x) = \begin{cases} 0 & ; x < 2,75 \\ \frac{x - 2,75}{(3,25 - 2,75)} & ; 2,75 \leq x \leq 3,25 \\ 1 & ; x > 3,25 \end{cases}$$

Himpunan *fuzzy* untuk setiap atribut panjang (m) tebu menggunakan kurva berbentuk trapesium seperti pada Gambar 2.4.



Gambar 2.4 Himpunan Fuzzy Atribut Panjang Tebu

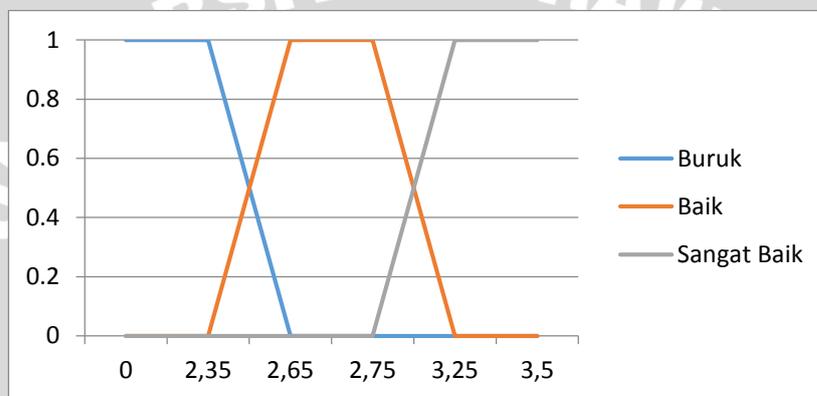
Atribut diameter (cm) tebu memiliki *range*, yaitu buruk (diameter < 2,5), baik (2,5 ≤ diameter < 3) dan sangat baik (diameter ≥ 3). Berikut merupakan himpunan *fuzzy* atribut diameter (cm) tebu:

$$\mu_{buruk}(x) = \begin{cases} 1 & ; x < 2,35 \\ \frac{2,65 - x}{(2,65 - 2,35)} & ; 2,35 \leq x \leq 2,65 \\ 0 & ; x > 2,65 \end{cases}$$

$$\mu_{baik}(x) = \begin{cases} 0 & ; x < 2,35 \vee x > 3,25 \\ \frac{x - 2,35}{(2,65 - 2,35)} & ; 2,35 \leq x \leq 2,65 \\ 1 & ; 2,65 < x < 2,75 \\ \frac{3,25 - x}{(3,25 - 2,75)} & ; 2,75 \leq x \leq 3,25 \end{cases}$$

$$\mu_{sangat\ baik}(x) = \begin{cases} 0 & ; x < 2,75 \\ \frac{x - 2,75}{(3,25 - 2,75)} & ; 2,75 \leq x \leq 3,25 \\ 1 & ; x > 3,25 \end{cases}$$

Himpunan *fuzzy* untuk setiap diameter (cm) tebu menggunakan kurva berbentuk trapesium seperti pada Gambar 2.5.



Gambar 2.5 Himpunan Fuzzy Atribut Diameter Tebu

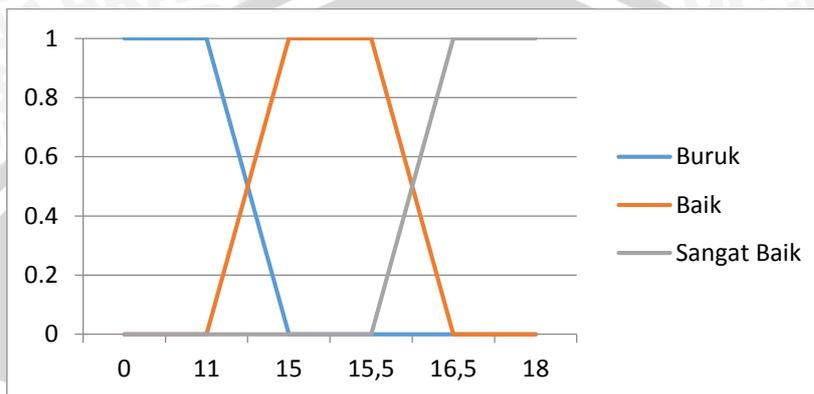
Atribut banyak ruas batang tebu dibagi menjadi 3 kelompok yaitu buruk (banyak ruas batang < 13), baik (13 ≤ banyak ruas batang < 16) dan sangat baik (banyak ruas batang ≥ 16). Berikut merupakan himpunan *fuzzy* atribut banyak ruas per batang tebu:

$$\mu_{buruk}(x) = \begin{cases} 1 & ; x < 11 \\ \frac{15 - x}{(15 - 11)} & ; 11 \leq x \leq 15 \\ 0 & ; x > 15 \end{cases}$$

$$\mu_{baik}(x) = \begin{cases} 0 & ; x < 11 \vee x > 16,5 \\ \frac{x - 11}{(15 - 11)} & ; 11 \leq x \leq 15 \\ 1 & ; 15 < x < 15,5 \\ \frac{16,5 - x}{(16,5 - 15,5)} & ; 15,5 \leq x \leq 16,5 \end{cases}$$

$$\mu_{sangat\ baik}(x) = \begin{cases} 0 & ; \quad x < 15,5 \\ \frac{x - 15,5}{(16,5 - 15,5)} & ; \quad 15,5 \leq x \leq 16,5 \\ 1 & ; \quad x > 16,5 \end{cases}$$

Himpunan *fuzzy* untuk setiap atribut banyak ruas per batang tebu menggunakan kurva berbentuk trapesium seperti pada Gambar 2.6.



Gambar 2.6. Himpunan Fuzzy Atribut Banyak Ruas per Batang Tebu

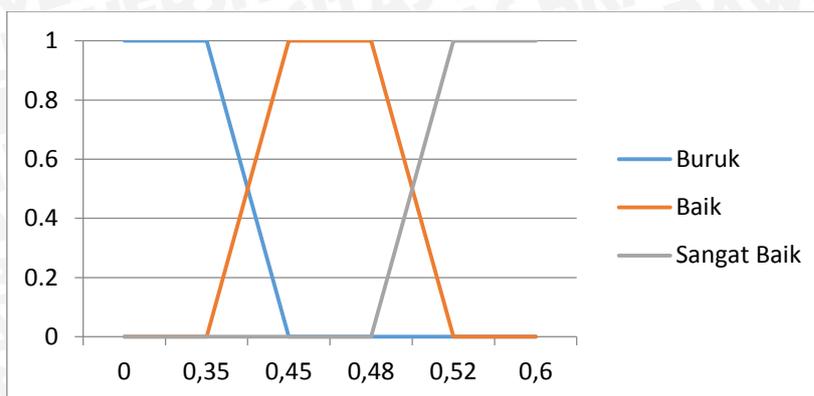
Atribut berat per meter tebu dibagi menjadi 3 kelompok yaitu buruk (berat < 0,4), baik (0,4 <= berat < 0,5) dan sangat baik (berat >= 0,5). Berikut merupakan himpunan *fuzzy* atribut berat per meter:

$$\mu_{buruk}(x) = \begin{cases} 1 & ; \quad x < 0,35 \\ \frac{0,45 - x}{(0,45 - 0,35)} & ; \quad 0,35 \leq x \leq 0,45 \\ 0 & ; \quad x > 0,45 \end{cases}$$

$$\mu_{baik}(x) = \begin{cases} 0 & ; \quad x < 0,35 \vee x > 0,52 \\ \frac{x - 0,35}{(0,45 - 0,35)} & ; \quad 0,35 \leq x \leq 0,45 \\ 1 & ; \quad 0,45 < x < 0,48 \\ \frac{0,52 - x}{(0,52 - 0,48)} & ; \quad 0,48 \leq x \leq 0,52 \end{cases}$$

$$\mu_{sangat\ baik}(x) = \begin{cases} 0 & ; \quad x < 0,48 \\ \frac{x - 0,48}{(0,52 - 0,48)} & ; \quad 0,48 \leq x \leq 0,52 \\ 1 & ; \quad x > 0,52 \end{cases}$$

Himpunan *fuzzy* untuk setiap atribut berat per meter menggunakan kurva berbentuk trapesium seperti pada Gambar 2.7.



Gambar 2.7. Himpunan Fuzzy Atribut Berat Per Meter

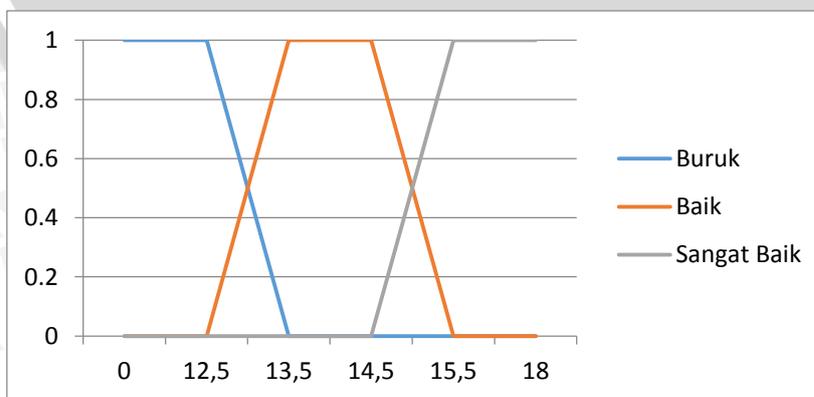
Atribut persentase (%) brix dibagi menjadi 3 kelompok yaitu buruk ($brix < 13$), baik ($13 \leq brix < 15$) dan sangat baik ($brix \geq 15$). Berikut merupakan himpunan *fuzzy* atribut persentase (%) brix:

$$\mu_{buruk}(x) = \begin{cases} 1 & ; \quad x < 12,5 \\ \frac{13,5 - x}{(13,5 - 12,5)} & ; \quad 12,5 \leq x \leq 13,5 \\ 0 & ; \quad x > 13,5 \end{cases}$$

$$\mu_{baik}(x) = \begin{cases} 0 & ; \quad x < 12,5 \vee x > 15,5 \\ \frac{x - 12,5}{(13,5 - 12,5)} & ; \quad 12,5 \leq x \leq 13,5 \\ \frac{15,5 - x}{(15,5 - 14,5)} & ; \quad 14,5 \leq x \leq 15,5 \end{cases}$$

$$\mu_{sangat\ baik}(x) = \begin{cases} 0 & ; \quad x < 14,5 \\ \frac{x - 14,5}{(15,5 - 14,5)} & ; \quad 14,5 \leq x \leq 15,5 \\ 1 & ; \quad x > 15,5 \end{cases}$$

Himpunan *fuzzy* untuk setiap atribut persentase (%) brix menggunakan kurva berbentuk trapesium seperti pada Gambar 2.8.



Gambar 2.8. Himpunan Fuzzy Atribut Persentase Brix

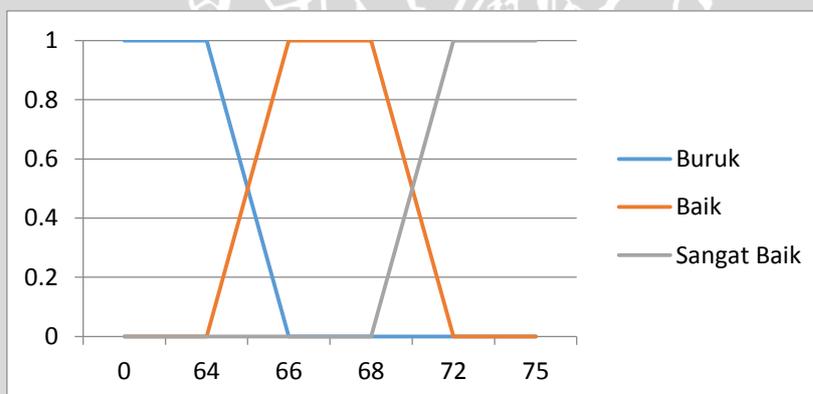
Atribut harkat kemurnian (HK) dibagi menjadi 3 kelompok yaitu buruk ($HK < 65$), baik ($65 \leq HK < 70$) dan sangat baik ($HK \geq 70$). Berikut merupakan himpunan *fuzzy* atribut harkat kemurnian :

$$\mu_{buruk}(x) = \begin{cases} 1 & ; x < 64 \\ \frac{66-x}{(66-64)} & ; 64 \leq x \leq 66 \\ 0 & ; x > 66 \end{cases}$$

$$\mu_{baik}(x) = \begin{cases} 0 & ; x < 64 \vee x > 72 \\ \frac{x-64}{(66-64)} & ; 64 \leq x \leq 66 \\ 1 & ; 66 < x < 68 \\ \frac{72-x}{(72-68)} & ; 68 \leq x \leq 72 \end{cases}$$

$$\mu_{sangat\ baik}(x) = \begin{cases} 0 & ; x < 68 \\ \frac{x-68}{(72-68)} & ; 68 \leq x \leq 72 \\ 1 & ; x > 72 \end{cases}$$

Himpunan *fuzzy* untuk setiap atribut harkat kemurnian menggunakan kurva berbentuk trapesium seperti pada Gambar 2.9.



Gambar 2.9. Himpunan Fuzzy Atribut Harkat Kemurnian

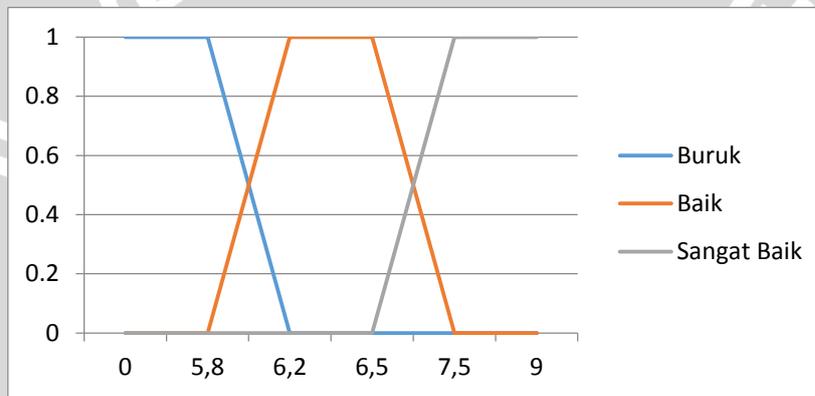
Atribut nira dibagi menjadi 3 kelompok yaitu buruk ($nira < 6$), baik ($6 \leq nira < 7$) dan sangat baik ($nira \geq 7$). Berikut merupakan himpunan *fuzzy* atribut nira :

$$\mu_{buruk}(x) = \begin{cases} 1 & ; x < 5,8 \\ \frac{6,2-x}{(6,2-5,8)} & ; 5,8 \leq x \leq 6,2 \\ 0 & ; x > 6,2 \end{cases}$$

$$\mu_{baik}(x) = \begin{cases} 0 & ; x < 5,8 \vee x > 7,5 \\ \frac{x - 5,8}{(6,2 - 5,8)} & ; 5,8 \leq x \leq 6,2 \\ 1 & ; 6,2 < x < 6,5 \\ \frac{7,5 - x}{(7,5 - 6,5)} & ; 6,5 \leq x \leq 7,5 \end{cases}$$

$$\mu_{sangat\ baik}(x) = \begin{cases} 0 & ; x < 6,5 \\ \frac{x - 6,5}{(7,5 - 6,5)} & ; 6,5 \leq x \leq 7,5 \\ 1 & ; x > 7,5 \end{cases}$$

Himpunan *fuzzy* untuk setiap atribut nira menggunakan kurva berbentuk trapesium seperti pada Gambar 2.10.



Gambar 2.10. Himpunan Fuzzy Atribut Nira

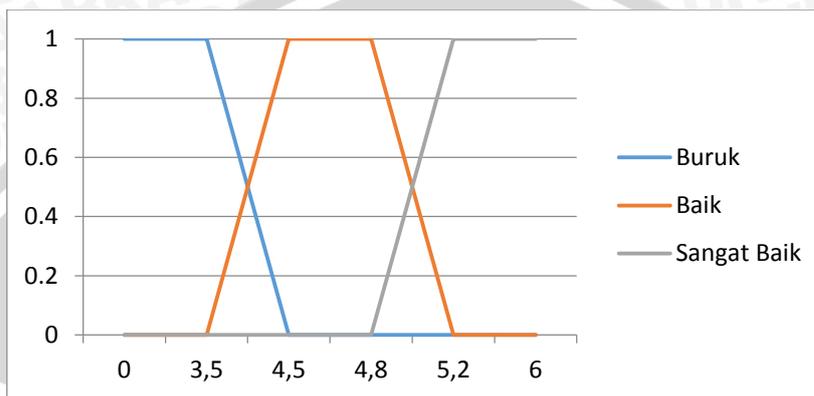
Atribut rendemen dibagi menjadi 3 kelompok yaitu buruk (rendemen<4), baik (4<=rendemen<5) dan sangat baik (rendemen>=5). Berikut merupakan himpunan *fuzzy* atribut rendemen :

$$\mu_{buruk}(x) = \begin{cases} 1 & ; x < 3,5 \\ \frac{4,5 - x}{(4,5 - 3,5)} & ; 3,5 \leq x \leq 4,5 \\ 0 & ; x > 4,5 \end{cases}$$

$$\mu_{baik}(x) = \begin{cases} 0 & ; x < 3,5 \vee x > 5,2 \\ \frac{x - 3,5}{(4,5 - 3,5)} & ; 3,5 \leq x \leq 4,5 \\ 1 & ; 4,5 < x < 4,8 \\ \frac{5,2 - x}{(5,2 - 4,8)} & ; 4,8 \leq x \leq 5,2 \end{cases}$$

$$\mu_{\text{sangat baik}}(x) = \begin{cases} 0 & ; x < 4,8 \\ \frac{x - 4,8}{(5,2 - 4,8)} & ; 4,8 \leq x \leq 5,2 \\ 1 & ; x > 5,2 \end{cases}$$

Himpunan *fuzzy* untuk setiap atribut rendemen menggunakan kurva berbentuk trapesium seperti pada Gambar 2.11.



Gambar 2.11. Himpunan Fuzzy Atribut Rendemen

Atribut klasifikasi selanjutnya akan disebut sebagai CLASS, direpresentasikan oleh empat peubah linguistik, yaitu “BL”, “HW”, “PS.881” dan “PS.862”. Keempat nilai linguistik tersebut didefinisikan sebagai berikut :

$$“BL” = 1; “HW” = 2; “PS.881” = 3; “PS.862” = 4$$

2.8. Akurasi

Data uji digunakan untuk memperkirakan keakuratan aturan klasifikasi. Jika akurasi dianggap dapat diterima, aturan dapat diterapkan terhadap klasifikasi data baru. Keakuratan model pada satu data uji yang diberikan adalah persentase sampel data uji yang diklasifikasikan dengan benar oleh model. Untuk setiap sampel uji, kelas dikenal dengan membandingkan prediksi kelas model pembelajaran pada sampel [HAK-01]. Pengenalan akurasi dari sistem diperkirakan sesuai dengan persamaan berikut [MHS-12]:

$$\text{Recognition accuracy (\%)} = \frac{\text{correct recognition}}{\text{total Number of testing sample for each digit}} \times 100\% \quad (2-10)$$

