

PREDIKSI *CUSTOMER CHURN*
MENGUNAKAN ALGORITMA *FUZZY ITERATIVE*
DICHOTOMISER 3

Meyrina Herawati, Imam Mukhlash, Inu L. Wibowo

Jurusan Matematika, FMIPA ITS Surabaya

meyrina.herawati@yahoo.com, imammukhlash@gmail.com

Abstrak

Setiap industri telekomunikasi sering menawarkan layanan yang menggiurkan. Hal ini bertujuan untuk menarik pelanggan sebanyak mungkin agar pendapatan yang masuk semakin optimal. Akan tetapi hal-hal seperti ini yang menyebabkan fenomena churn terjadi, dimana pelanggan memutuskan berhenti berlangganan dan berpindah dari satu provider ke provider lainnya. Tiap industri telekomunikasi juga harus memiliki program retaining yang efektif dan efisien, terlebih dahulu perusahaan memprediksi pelanggan yang berpotensi churn, kemudian menjadikan pelanggan tersebut sebagai fokus usaha retention. Penelitian ini bertujuan untuk melakukan pengembangan model churn prediction (classification) dengan menggunakan fuzzy decision tree. Fuzzy decision tree merupakan salah satu metode klasifikasi berbasis pohon keputusan dengan memasukkan konsep himpunan fuzzy dalam proses pembentukannya. Setelah tahap implementasi, hasil dari penelitian ini berupa model pohon keputusan prediksi customer churn. Dari berbagai uji coba didapat nilai akurasi terbaik 87

Kata kunci : *Data Mining, Fuzzy Decision Tree, Fuzzy Logic, Management Churn.*

1. Pendahuluan

Setiap perusahaan memiliki cara tersendiri dalam menawarkan layanan yang berkualitas dengan harga seminimum mungkin. Hal ini diharapkan dapat menarik pelanggan sebanyak mungkin agar pendapatan yang masuk ke perusahaan semakin optimal pula. Hal ini menyebabkan fenomena *churn* terjadi, dimana para pelanggan memutuskan untuk pindah dari satu *provider* ke *provider* yang lain karena tergiur penawaran fasilitas dan harga yang lebih menarik. Fenomena *churn* ini tentu saja meresahkan, karena jika tidak dicegah dan ditangani akan berakibat pada penurunan *revenue* perusahaan [1].

Bagi banyak perusahaan, menemukan alasan kehilangan pelanggan, mengukur loyalitas pelanggan dan mendapatkan kembali pelanggan menjadi konsep yang sangat penting untuk mencegah *churners*. Data yang terkait dengan sektor telekomunikasi meningkat dengan sangat cepat karena teknologi yang selalu berkembang, peningkatan jumlah pelanggan dan bertambahnya jenis layanan. Ekspansi yang tidak terkendali dan sangat cepat di bidang ini menyebabkan bertambahnya kerugian yang dialami oleh perusahaan. Oleh karena itu, perkembangan metode analisis baru telah menjadi suatu keharusan. Hal ini telah menjadi pelopor dalam berbagai penelitian di sektor telekomunikasi yang mengalami kerugian besar.

Data mining sebagai salah satu alat analisis pada *analytical CRM* telah terbukti kegunaannya bagi suatu perusahaan dalam memprediksi pelanggan yang akan meninggalkan perusahaan, kapan mereka akan *churn*, dan tingkat prediksi tersebut [2]. Dengan memprediksi jumlah pelanggan yang akan mengganti jasa pelayanan secara terus-menerus, perusahaan dapat menciptakan promosi yang bertujuan untuk meningkatkan loyalitas pelanggan dan meningkatkan strategi penjualan untuk mendapatkan pelanggan baru dan melakukan retensi pelanggan.

Metode-metode data mining yang pernah digunakan untuk analisis *churn* antara lain *Bayesian Belief Network* [3], *Decision Tree* [4], *Artificial Neural Network* [5], *Cross Industry Standard Process-Data Mining (CRISP-DM)* [6], *Nave Bayes Classifier* [6] dan masih banyak lagi.

Salah satu penelitian mengenai *churn management* dilakukan oleh Zakki Madawvossi menyimpulkan bahwa model pendekatan *decision tree* yang memiliki performansi paling tinggi adalah model DT (C4.5 - kategorikal)[5]. Pada penelitian ini digunakan metode *Fuzzy Decision Tree*. *Fuzzy Decision tree* merupakan penggabungan data mining dengan *fuzzy logic*. Data mining yang digunakan dalam penelitian ini adalah *Decision Tree*. *Decision Tree* merupakan metode yang terkenal daripada metode lainnya karena metode ini tidak membutuhkan pengetahuan yang lebih atau pengaturan parameter [7]. Penggabungan metode *decision tree* dengan *fuzzy* memungkinkan untuk menggunakan nilai-nilai numerik yang dihubungkan dengan atribut kuantitatif yang masing-masing memiliki nilai derajat

keanggotaan. Penggunaan teknik *fuzzy* dalam *decision tree* dapat meningkatkan kemampuan atribut-atribut kuantitatif dalam melakukan penggolongan pada saat pelatihan. Selain itu, proses pengujian menggunakan metode *Fuzzy Inference Sytem Mamdani* untuk mendapatkan output yang baik berdasarkan nilai derajat keanggotaan masing-masing atribut [8].

Pada penelitian ini, algoritma yang digunakan dalam *fuzzy decision tree* adalah ID3. Algoritma ini menggunakan teori informasi untuk menentukan atribut mana yang paling informatif. Namun ID3 sangat tidak stabil dalam melakukan klasifikasi yang berkenaan dengan gangguan kecil pada data pelatihan. Logika *fuzzy* dapat memberikan suatu peningkatan dalam melakukan klasifikasi pada saat pelatihan. Metode *fuzzy decision tree* menggunakan algoritma ID3 ini telah dilakukan pada beberapa riset antara lain riset yang dilakukan F. Romansyah tahun 2009 pada data diabetes [9]. Berdasarkan beberapa penelitian yang telah dilakukan sebelumnya, maka pada penelitian ini dilakukan penelitian mengenai prediksi *customer churn* dengan menggunakan algoritma *fuzzy decision tree*.

2. Fuzzy Decision Tree dan Aplikasinya

2.1. Fuzzy Decision Tree

Fuzzy decision tree merupakan suatu pendekatan yang populer dan praktis dalam *machine learning* untuk menyelesaikan permasalahan klasifikasi yang memuat ketidakpastian. *Fuzzy decision tree* memungkinkan untuk menggunakan nilai-nilai *numeric-symbolic* selama konstruksi atau saat mengklasifikasikan kasus-kasus baru. Salah satu algoritma yang dapat digunakan untuk membangun *fuzzy decision tree* adalah algoritma *Iterative Dichotomiser 3* (ID3).

Algoritma *fuzzy ID3* merupakan algoritma yang efisien untuk membuat suatu *fuzzy decision tree*. Algoritma *fuzzy ID3* adalah sebagai berikut [3][9]:

1. Create a root node that has a set of fuzzy data with membership value 1.
2. If a node t with a fuzzy set of data D satisfies the following conditions, then it is a leaf node and assigned by the class name
 - The proportion of class C_k is greater than or equal to θ_r .
 - The number of a data set is less than θ_n
 - There are no attributes for more classifications
3. If a node D does not satisfy the above conditions, then it is not a leaf-node. And a new sub-node is generated as follows:

- For $A'_i (i = 1, \dots, L)$ calculate the information gain, and select the test attribute A_{max} that maximizes them.
- Devide D into fuzzy subset D_1, \dots, D_n according to A_{max} , where the membership value of the data in D_j is the product of the membership value in D and the value of $F_{max,j}$ of the value of A_{max} in D .
- Generate new node t_1, \dots, t_n for fuzzy subsets D_1, \dots, D_n and label the fuzzy set $F_{max,j}$ to edges that connect between the nodes t_j and t .
- Replace D by $D_j (j = 1, 2, \dots, m)$ and repeat from 2 recursively)

Information gain adalah suatu nilai statistik yang digunakan untuk memilih atribut yang akan mengekspansi tree dan menghasilkan node baru pada algoritma ID3. Suatu *entropy* dipergunakan untuk mendefinisikan nilai *information gain*. *Entropy* dirumuskan sebagai berikut [3][9]:

$$H_s(S) = \sum_i^N - P_i * \log_2(P_i) \quad (1)$$

dengan P_i adalah rasio dari kelas C_i pada himpunan, contoh $S = \{x_1, x_2, \dots, x_k\}$

$$P_i = \frac{\sum_{j=1}^k x_j \in C_i}{S} \quad (2)$$

Information gain digunakan sebagai ukuran seleksi atribut, yang merupakan hasil pengurangan entropy dari himpunan contoh setelah membagi ukuran himpunan contoh dengan atributnya. *Information gain* didefinisikan sebagai berikut [3][9]:

$$G(S, A) = H(S) - \sum_{v \in Values(A)} \frac{|S_v|}{|S|} H(S_v) \quad (3)$$

dengan bobot $W_i = \frac{|S_v|}{|S|}$ adalah rasio dari data dengan atribut v pada himpunan contoh. Pada himpunan data *fuzzy*, terdapat penyesuaian rumus untuk menghitung nilai *entropy* untuk atribut dan *information gain* karena adanya ekspresi data *fuzzy*. Berikut adalah persamaan untuk mencari nilai *fuzzy entropy* dari keseluruhan data [3][9]:

$$H_f(S) = H_s(S) = \sum_i^N - P_i * \log_2(P_i) \quad (4)$$

Untuk menentukan *fuzzy entropy* dan *information gain* dari suatu atribut pada algoritma *fuzzy ID3* (FID3) digunakan persamaan sebagai berikut [3][9]:

$$H_f(S, A) = -\sum_{i=1}^c \frac{\sum_j^N \mu_{ij}}{S} \log_2 \frac{\sum_j^N \mu_{ij}}{S} \quad (5)$$

$$G_f(S) = H_f(S) - \sum_{v \subseteq A} \frac{|S_v|}{|S|} * H_f(S_v, A) \quad (6)$$

Dengan μ_j adalah nilai keanggotaan dari pola ke- j untuk kelas ke- i . $H_f(S)$ menunjukkan entropy dari himpunan S dari data pelatihan pada node. $|S_v|$ adalah ukuran dari subset $S_v \subseteq S$ dari data pelatihan x_j dengan v . $|S|$ menunjukkan ukuran dari himpunan S .

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 meningkatkan akurasi proses learning harus dihentikan lebih awal atau melakukan pemotongan *tree* secara umum. Untuk itu diberikan 2 (dua) *threshold* yang harus terpenuhi jika *tree* akan diekspansi, yaitu [3][9]:

- *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*.
- *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*.

2.2. *Customer Churn* di Bidang Telekomunikasi

Masalah terbesar yang sedang dihadapi industri telekomunikasi saat ini adalah fenomena yang menunjukkan perpindahan/pergerakan pelanggan dari satu *provider* ke *provider* lainnya. Fenomena ini yang dikenal dengan istilah *customer churn* [10].

Customer churn dapat disebabkan oleh banyak hal, mulai dari tarif yang kompetitif antar operator, fitur dan fasilitas yang kompetitif, sampai bagaimana provider melayani, berinteraksi, dan mengelola hubungannya dengan pelanggan-pelanggannya.

Customer churn di industri telekomunikasi merupakan masalah yang sangat besar. Jika pada industri lain, misalnya industri kartu kredit, nilai *customer churn* yang terjadi sekitar 0.4% tiap bulan, di industri telekomunikasi dapat terjadi sampai 2,2% per bulan. Dapat dikatakan bahwa sekitar 27% pelanggan *provider churn* setiap tahunnya [11].

Masalah *customer churn* ini menjadi krusial, karena biaya yang dikeluarkan untuk mendapatkan pelanggan baru, untuk iklan, marketing, komisi, dan lain-lain

akan jauh lebih besar dibandingkan biaya yang harus dikeluarkan untuk menjaga pelanggan yang sudah ada. Ditambah lagi belum kebanyakan pelanggan baru cenderung tidak lebih menghasilkan keuntungan dibandingkan pelanggan yang sudah lama dan bertahan [12].

3. Perancangan Sistem

3.1. Perancangan Data

Sesuai dengan permasalahan yang diangkat dalam penelitian ini, terdapat lima atribut kondisional data yang akan digunakan, yaitu *months*, *totmou*, *totcalls*, *custcare*, dan *phones*. Sedangkan untuk atribut keputusannya terdapat dua keputusan, yaitu *churn* dan tidak *churn*. Data training yang digunakan untuk pembentukan aturan dan tree ada 3 macam, yaitu data training yang berjumlah 400, 500 dan 700, sedangkan data yang digunakan untuk proses pengujian hanya terdapat satu macam yaitu data uji yang berjumlah 100.

Kelima atribut tersebut di transformasi kedalam bentuk *fuzzy* dengan menggunakan 2 fungsi yaitu kurva trapesium dan kurva segitiga. Data Luaran yang dihasilkan dari proses prediksi *customer churn* dengan menggunakan metode FID3 berupa aturan dan hasil ujicoba dari model FID3.

3.2. Gambaran Sistem Secara Umum

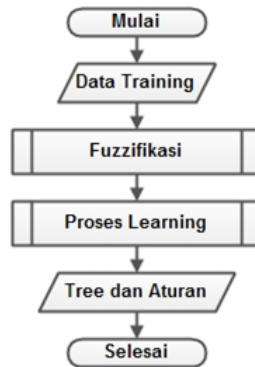
Gambaran sistem secara umum proses prediksi customer churn dengan menggunakan *Fuzzy Iterative Dichotomiser 3* (FID3) dalam bahasan ini terdapat 2 proses, yaitu proses pembentukan aturan dan tree serta proses pengujian model *classifier*.

3.2.1. Pembentukan Aturan dan Tree

Pada proses pembentukan aturan dan tree terdiri dari beberapa tahapan, antara lain yaitu proses fuzzifikasi data, proses *learning*. Alur tahapan proses pembentukan aturan adalah sebagai berikut:

1. Masukan berupa data training yaitu data pelanggan yang terdiri dari beberapa atribut.
2. Data tersebut diubah ke dalam bentuk data *fuzzy*.
3. Data training yang berbentuk data *fuzzy* masuk ke dalam proses *learning* untuk diklasifikasikan berbentuk tree. Dalam proses *learning* diterapkan algoritma *fuzzy iterative dichotomiser 3*. Selanjutnya dilakukan pembacaan tree untuk mendapatkan keluaran berupa aturan.

Diagram alur sistem proses pembentukan aturan pada Gambar 1:



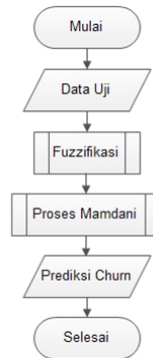
Gambar 1: Alur Proses Pembentukan Aturan

3.2.2. Pembentukan Aturan dan Tree

Proses Pengujian merupakan proses yang dilakukan untuk menguji data prediksi churn ke dalam aturan yang telah terbentuk. Dalam proses ini juga dilakukan penghitungan keakurasian aturan yang terbentuk dari proses sebelumnya, yaitu proses pembentukan aturan dan tree dengan menguji beberapa data set untuk diklasifikasi menggunakan inferensi fuzzy sistem Mamdani. Langkah-langkah proses pengujian aturan sebagai berikut:

1. Masukan merupakan data uji dari prediksi churn yang meliputi 5 atribut.
2. Data uji tersebut dilakukan proses fuzzifikasi untuk memperoleh derajat keanggotaan untuk masing-masing atribut
3. Kemudian dilakukan proses inferensi menggunakan metode Mamdani dan menggunakan aturan yang telah terbentuk sebelumnya.
4. Hasil dari inferensi mamdani tersebut berupa Status pelanggan yang merupakan kelas dari pelanggan tersebut

Diagram alur proses pengujian aturan dapat dilihat pada Gambar 2



Gambar 2: Alur Proses Pengujian

4. Hasil Pengujian Sistem

4.1. Lingkungan Pengujian Sistem

Perangkat yang digunakan dalam pengujian sistem terdiri dari beberapa perangkat keras dan perangkat lunak. Perangkat keras yang digunakan yaitu komputer dengan Prosesor Intel(R) Core(TM) i3-2630 QM CPU@2.00 GHz 2.00 GHz, RAM 6GB, dan Hard Disk 250GB. Sedangkan perangkat lunak yang digunakan adalah notepad++.

4.2. Analisis Hasil Pengujian

Terdapat 2 pengujian yang dilakukan yaitu:

1. Pengujian jumlah aturan yang terbentuk dari proses pembentukan aturan dan pohon (*tree*).
2. Pengujian nilai akurasi klasifikasi.

Uji coba ini dilakukan dengan menggunakan tiga macam data training dan satu macam data uji. Data training yang pertama berjumlah 400, yang kedua berjumlah 500 dan yang terakhir berjumlah 700, sedangkan untuk data uji yang digunakan berjumlah 100. Masing-masing data training dilakukan proses pembentukan aturan dan pohon dengan perubahan representasi kurva fuzzy dan memasukkan nilai kombinasi FCT dan LDT yang berbeda-beda. Nilai FCT yang diujikan adalah antara 50%, 60%, 70%, 80% dan 90% sedangkan nilai LDT yang diujikan adalah antara 3%, 8%, dan 15%. Dengan perubahan representasi kurva fuzzy serta memasukkan kombinasi nilai FCT dan LDT ini akan diperoleh jumlah *rule* atau aturan yang berbeda-beda. Pada pengujian tingkat akurasi, kelas

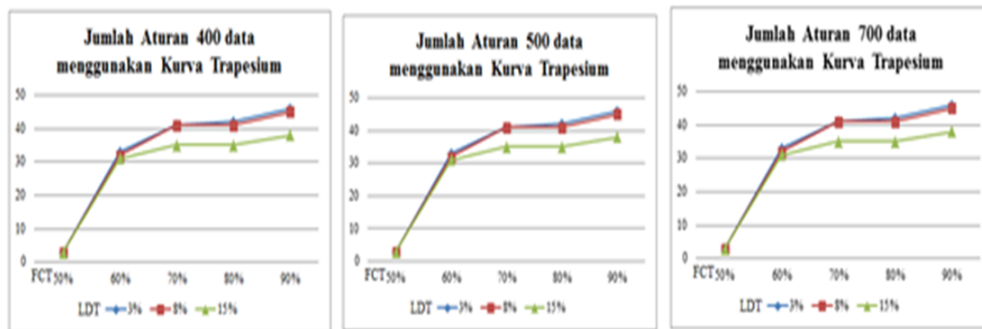
output yang dihasilkan dibandingkan dengan kelas output pada data asli. Untuk mendapatkan kelas output, dilakukan proses inferensi Mamdani yang disesuaikan dengan aturan-aturan yang terbentuk dalam proses sebelumnya.

Jum	400			500			700		
FCT	LDT								
	3%	8%	15%	3%	8%	15%	3%	8%	15%
50%	3	3	3	3	3	3	3	3	3
60%	28	27	26	33	32	30	33	32	31
70%	32	31	29	35	34	32	41	41	35
80%	36	35	33	39	38	34	46	45	38
90%	36	35	33	39	38	34	46	45	38

Tabel 1: Jumlah Aturan yang Terbentuk Menggunakan Kurva Trapesium

Jum	400			500			700		
FCT	LDT								
	3%	8%	15%	3%	8%	15%	3%	8%	15%
50%	3	3	3	3	3	3	3	3	3
60%	26	25	20	28	26	21	28	26	23
70%	30	27	22	31	28	23	32	28	23
80%	30	27	22	31	28	23	32	28	28
90%	30	27	22	31	28	23	32	28	28

Tabel 2: Jumlah Aturan yang Terbentuk menggunakan Kurva Segitiga



Gambar 3: Grafik Jumlah Aturan yang Terbentuk menggunakan Kurva Trapesium



Gambar 4: Grafik Jumlah Aturan yang Terbentuk dengan Menggunakan Kurva Segitiga

Pada Tabel 1 dan Tabel 2 dapat dilihat jumlah aturan yang terbentuk dari proses pembentukan aturan dan tree, jumlah aturan yang terbentuk tersebut bervariasi seiring bergantinya FCT, LDT, representasi kurva fuzzy dan jumlah data training yang digunakan. Dari Grafik 1 dan Grafik 2, dapat dilihat bahwa semakin tinggi nilai FCT akan menyebabkan jumlah aturan yang dihasilkan semakin banyak, jumlah aturan juga dipengaruhi oleh nilai LDT, semakin rendah nilai LDT semakin banyak jumlah aturan yang terbentuk.

Jum	400			500			700		
FCT	LDT								
	3%	8%	15%	3%	8%	15%	3%	8%	15%
50%	82%	82%	82%	87%	87%	87%	87%	87%	87%
60%	86%	86%	86%	87%	87%	87%	82%	82%	82%
70%	87%	87%	87%	86%	86%	86%	82%	82%	82%
80%	87%	87%	87%	86%	86%	86%	82%	82%	82%
90%	87%	87%	87%	86%	86%	86%	82%	82%	82%

Tabel 3: Nilai Akurasi Menggunakan Kurva Trapesium

Pada Tabel 3 dan Tabel 4 dapat dilihat nilai akurasi aturan yang dihasilkan pada uji coba ini, dipengaruhi oleh perubahan nilai FCT, nilai LDT, representasi kurva fuzzy dan jumlah data training yang digunakan. Dari Grafik 3 untuk data training yang berjumlah 400, semakin rendah nilai FCT yang digunakan semakin rendah juga nilai akurasinya sedangkan untuk data training yang berjumlah 500 dan 700, semakin rendah nilai FCT semakin tinggi nilai akurasinya. Pada Grafik 4, dapat dilihat semakin tinggi nilai FCT semakin tinggi nilai akurasi yang dihasilkan. Selain nilai FCT yang berpengaruh dengan nilai akurasi, nilai LDT juga ikut mempengaruhi nilai akurasi dari sebuah aturan. Semakin tinggi nilai LDT yang digunakan, semakin rendah nilai akurasi dari aturan yang terbentuk.

Jum	400			500			700		
FCT	LDT								
	3%	8%	15%	3%	8%	15%	3%	8%	15%
50%	71%	71%	71%	81%	81%	81%	80%	80%	80%
60%	84%	84%	87%	87%	85%	85%	86%	84%	83%
70%	84%	84%	87%	87%	87%	85%	86%	84%	83%
80%	84%	84%	87%	87%	87%	85%	86%	84%	83%
90%	84%	84%	87%	87%	87%	85%	86%	84%	83%

Tabel 4: Nilai Akurasi Menggunakan Kurva Segitiga



Gambar 5: Grafik Nilai Akurasi Menggunakan Kurva Trapesium



Gambar 6: Grafik 4 Grafik Nilai Akurasi Menggunakan Kurva Segitiga

Dari hasil uji coba ini dapat dilihat pada grafik yang menggunakan representasi kurva trapesium mengalami *overfitting* karena semakin rendah nilai FCT, semakin tinggi nilai akurasi aturan yang terbentuk. Pada penelitian yang dilakukan oleh Romansyah [9] juga terjadi hal yang sama yaitu *overfitting*. Hal tersebut terjadi karena tingginya nilai FCT yang digunakan dalam proses pembentukan mengakibatkan tree akan terus di ekspansi maka tree memiliki node-node yang mengandung data yang mengalami kesalahan klasifikasi.

Dari keseluruhan uji coba representasi kurva fuzzy, nilai FCT, nilai LDT dan jumlah data training berpengaruh terhadap jumlah aturan yang terbentuk dan tingkat keakurasian aturan yang terbentuk dari proses pembentukan aturan dan tree.

Dari uji coba ini representasi kurva trapesium lebih baik dengan menggunakan data training 700 diperoleh nilai akurasi adalah 87% terdapat pada kombinasi FCT 50% dan LDT 3% , 8% dan 15% serta aturan yang terbentuk sebanyak 3 aturan. Aturan terbanyak yang terbentuk adalah 46 aturan terdapat pada data training yang berjumlah 700 dengan nilai FCT 90% dan nilai LDT 3% dan juga menggunakan representasi kurva trapesium.

5. Kesimpulan

Pada Penelitian ini telah dibangun perangkat lunak untuk prediksi customer churn menggunakan *Fuzzy Iterative Dichotomiser 3*. Parameter-parameter yang digunakan untuk prediksi *customer churn* yaitu lama bulan penggunaan, total panggilan, total menit panggilan, total panggilan *customer service*, dan jumlah telepon. Berdasarkan analisis terhadap hasil pengujian perangkat lunak untuk prediksi *customer churn* pada dataset maka dapat diambil beberapa kesimpulan sebagai berikut:

1. Metode *Fuzzy Iterative Dichotomiser 3* (FID3) dapat diimplementasikan untuk prediksi *customer churn*. Proses utama yang dilakukan, yaitu proses pembentukan aturan dan tree dengan menggunakan metode Fuzzy ID3 terhadap data training dan proses pengujian dengan menggunakan metode *Fuzzy Inference System Mamdani* berdasarkan aturan yang terbentuk dari proses pembentukan aturan dan pohon. Hasil dari proses pengujian menggunakan metode *Fuzzy Inference System Mamdani* selanjutnya digunakan untuk menentukan kelas klasifikasi.
2. Dari berbagai uji coba didapat nilai akurasi terbaik 87

Pustaka

- [1] SHAN JIN, YUN MENG, CHUNFEN FAN, FENG PENG, QINGZHANG CHEN, *The Research on Applying Data Mining to Telecom Churn Management*, 2008.
- [2] DYCHE, JILL, *The CRM Handbook*, Addison-Wesley, 2002.
- [3] SUHARSO, WILDAN. DJUNAIDY, ARIF, *Analisis Customer Churn menggunakan bayesian belief network(studi kasus: perusahaan layanan internet)*, Surabaya: Institut Teknologi Sepuluh Nopember, 2013.

-
- [4] SUHARTONO, ENTOT, *Prediksi perilaku loyalitas pelanggan tehnik data mining decision tree*, Stie Bank BPD jateng, 2011.
 - [5] MADAWWOSSI, ZAKKI, *Aplikasi Data Mining untuk Churn Prediction (Studi Kasus Pelanggan Flexi Classy PT. Telkom Divre I Sumatera)*, Surabaya: Institut Teknologi Sepuluh Nopember, 2009.
 - [6] ROSMALA, DEWI, *Implementasi Crips-DM dan Nave Bayes Classifier pada data mining churn prediction*, Bandung: Institut Teknologi Nasional, 2013.
 - [7] HAN, JIAWEI DAN KAMBER, MICHELINE, *Data Mining: Concepts and Techniques*, San Fransisco : Morgan Kaufmann Publishers, 2003.
 - [8] LIANG, G, *A Comparative Study of Three Decision Tree Algorithms: ID3, Fuzzy ID3 and Probalistic Fuzzy ID3*, Informatics & Economics Erasmus University Rotterdam, The Netherlands, 2005.
 - [9] ROMANSYAH, F., SITANGGANG, I. S., DAN NURDIATI, S., *Fuzzy Decision Tree dengan Algoritma ID3 pada Data Diabetes*, Internetworking Indonesia Journal, Vol. 1, No. 2, Special Issue on Data Mining, 2009.
 - [10] MUTANEN, TEEMU, *Customer Churn Analysis- a case study*, Technical Research Centre of Finland.Finland, 2006.
 - [11] BERSON, ALEX, DKK, *Building Data Mining Application for CRM*, McGraw-Hill, New York, 1999.
 - [12] VALENTINUS RAMA KURNIANGGA, TRI DANUR WENDA SUKARDJO, *Analisa dan Perancangan Data dan Text Mining Terhadap Customer Relationship Management System pada PT. Excelcomindo Pratama.Tbk. Jakarta*, 2009.
 - [13] PUSPITASARI, DIANTIKA, SETIAWAN, BUDI DARMA, DAN MARJI, *Implementasi Fuzzy Iterative Dichotomiser 3(FID3) dalam Klasifikasi Tebu*, Malang : Universitas Brawijaya, 2014.