

**KRITERIA NASABAH NON MUSLIM MENABUNG (PENGALIAN
DATA MENGGUNAKAN KLASIFIKASI ALGORITMA C4.5 _ STUDI
KASUS DI PT. BANK BRI SYARIAH
KANTOR CABANG MEDAN)**

Nurbaiti

Fakultas Ekonomi dan Bisnis Islam Universitas Islam Negeri Sumatera Utara
nurbaiti@uinsu.ac.id

Abstract

Algorithm is a classification algorithm C4.5 decision tree data type (decision tree). C4.5 decision tree algorithm is built in three phases, namely the selection of attributes as root, create a branch for each value and divide the cases in the branch. These stages will be repeated for each branch until all cases the branches have the same class. Of the completion of the decision tree will produce some rule or knowledge of a case. In this thesis, the author classifies the criteria of non-Muslim customers save money in Islamic banks. The resulting rule greatly determined by the input variables in the set such as gender, age, religion, education level, employment and decision variables, namely customer saving and not saving. Application of C4.5 algorithm can determine the criteria that non-Muslim customers saving in PT. Bank BRI Syariah Kantor Cabang Medan

Keywords: Data Mining, Classification, C4.5 algorithm, Decision Tree, Savings, Customer Criteria

Abstrak

Algoritma C4.5 adalah algoritma klasifikasi data bertipe pohon keputusan (*decision tree*). Pohon keputusan algoritma C4.5 dibangun dengan tiga tahap yaitu pemilihan atribut sebagai akar, membuat cabang untuk tiap-tiap nilai dan membagi kasus dalam cabang. Tahap-tahap ini akan diulangi untuk setiap cabang sampai semua kasus pada cabang memiliki kelas yang sama. Dari penyelesaian pohon keputusan maka akan di hasilkan beberapa *rule* atau pengetahuan dari suatu kasus. Dalam tesis ini penulis mengklasifikasikan kriteria nasabah non muslim menabung di bank syariah. Rule yang dihasilkan sangat di tentukan oleh variabel-variabel input yang di tetapkan antara lain jenis kelamin, umur, agama, pendidikan terakhir, pekerjaan dan variabel keputusannya nasabah menabung dan tidak menabung. Penerapan algoritma C4.5 dapat mengetahui kriteria nasabah non muslim yang menabung di PT. Bank BRI Syariah Kantor Cabang Medan

Kata Kunci: *Data Mining*, Klasifikasi, Algoritma C4.5, Pohon Keputusan, Tabungan, Kriteria Nasabah.

Pendahuluan

Dalam bidang perbankan, teknik *data mining* digunakan untuk mendukung cakupan yang luas dari aplikasi-aplikasi bisnis inteligen seperti *customer profiling*, *targeted marketing*, *workflow management*, *store layout* dan *fraud*

detection. Teknik *data mining* dapat digunakan untuk menjawab pertanyaan bisnis yang penting seperti ”Siapakan pelanggan yang akan paling banyak mendatangkan keuntungan ?” dan ” Seperti apa perkiraan pendapatan perusahaan tahun depan ?”.

Bank BRI Syariah, bank yang beroperasi berdasarkan system syariah, bank ini bukan hanya melayani nasabah muslim tapi juga non muslim. Untuk mengetahui kriteria nasabah non muslim yang menabung di bank syariah maka peneliti tertarik untuk melakukan sebuah penelitian yang berjudul “Kriteria Nasabah Non Muslim Menabung (Penggalian Data Dengan Menggunakan Metode Klasifikasi Algoritma C.45 _ Studi Kasus di PT. Bank BRI Syariah Kantor Cabang Medan”.

Landasan Teori

1. Data mining

Data mining merupakan bidang dari beberapa bidang keilmuan yang menyatukan teknik dari pembelajaran mesin, pengenalan pola, statistik, database dan visualisasi untuk penanganan permasalahan pengambilan informasi dari database yang besar (Larose, 2005).

Data mining adalah proses analisa data untuk menemukan suatu pola dari kumpulan data. *Data mining* mampu menganalisa data yang besar menjadi informasi berupa pola yang mempunyai arti bagi pendukung keputusan (Sunjana, 2010).

2. Knowledge Discovery in Database (KDD)

Data mining sering dianggap sebagai bagian dari *knowledge discovery in database* (KDD) yaitu sebuah proses mencari pengetahuan yang bermanfaat dari data, proses KDD secara garis besar dapat dijelaskan sebagai berikut (Sunjana, 2010)

a. Data Selection

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

b. Pre-processing/Cleaning

Sebelum proses *data mining* dapat dilaksanakan, perlu dilakukan proses *cleaning* pada data yang menjadi fokus KDD. Proses *cleaning* mencakup antara

lain membuang duplikasi data memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data, seperti kesalahan cetak.

c. *Transformation*

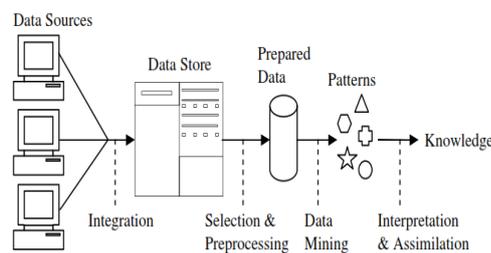
Coding adalah proses transformasi pada data yang telah dipilih, sehingga data tersebut sesuai untuk proses *data mining*. Proses *coding* dalam KDD merupakan proses kreatif dan sangat bergantung pada jenis atau pola informasi yang akan dicari dalam basis data.

d. *Data mining*

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

e. *Interpretation/Evaluation*

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



Gambar 2.1 Tahap-Tahap *Knowledge Discovery in Database*

3. Operasi *Data Mining*

Operasi *data mining* menurut sifatnya dibedakan menjadi dua, yaitu bersifat :

a. Prediksi (*prediction driven*)

Yaitu untuk menjawab pertanyaan apa dan sesuatu yang bersifat remang-remang atau transparan. Operasi prediksi digunakan untuk validasi hipotesis, *querying* dan pelaporan. Misalnya spreadsheet dan pivot table, analisis multidimensi (*dimensional summary*), OLAP (*Online Analytic Processing*) serta analisis statistic.

b. Penemuan (*discovery driven*)

Yaitu bersifat transparan dan untuk menjawab pertanyaan “mengapa?”. Operasi penemuan digunakan untuk analisis data eksplorasi, pemodelan prediktif, segmentasi database, analisis keterkaitan (*link analysis*) dan deteksi devias. (Fajar Astuti Hermawati, 2013).

4. Teknik *Data Mining*

Dalam prakteknya, dua tujuan utama dari *data mining* cenderung prediksi dan deskripsi. Prediksi melibatkan menggunakan beberapa variabel atau kolom dalam kumpulan data untuk memprediksi nilai yang tidak diketahui atau masa depan variabel lain yang menarik. Keterangan, di sisi lain, berfokus pada menemukan pola yang menggambarkan data yang dapat diinterpretasikan oleh manusia. Oleh karena itu, adalah mungkin untuk menempatkan kegiatan *data mining* menjadi satu dari dua kategori:

- a. *Data mining prediktif*, yang menghasilkan model sistem yang dijelaskan oleh himpunan data, atau
- b. *Data mining deskriptif*, yang menghasilkan baru, informasi trivial berdasarkan kumpulan data yang tersedia. (*Ebook – Data mining*)

Data mining dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan, yaitu (Larose, 2004):

1. Deskripsi

Terkadang peneliti dan analis secara sederhana ingin mencoba mencari cara untuk menggambarkan pola dan kecenderungan yang terdapat dalam data.

2. Estimasi

Estimasi hampir sama dengan klasifikasi, kecuali variabel target estimasi lebih ke arah numerik daripada ke arah kategori.

3. Prediksi

Prediksi hampir sama dengan klasifikasi dan estimasi, kecuali bahwa dalam prediksi nilai dari hasil akan ada di masa mendatang.

4. Klasifikasi

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

5. Pengklusteran

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

6. Asosiasi

Tugas asosiasi dalam *data mining* adalah menemukan atribut yang muncul dalam satu waktu. Dalam dunia bisnis lebih umum disebut analisis keranjang belanja.

5. Metode Klasifikasi (*Classification*)

Klasifikasi adalah suatu proses untuk menemukan satu model atau fungsi yang menjelaskan dan memisahkan *class* dari suatu data dengan tujuan agar dapat menggunakan model atau fungsi tersebut untuk memprediksi suatu *class* dari satu objek baru yang belum dikenali (Han & Kamber, 2006).

Menurut Dedy Rahman Wijaya (2005), *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. Model itu sendiri bisa berupa aturan jika-maka, berupa *decision tree*, formula matematis atau *neural network*.

6. Algoritma C4.5

Algoritma C4.5 adalah algoritma yang sudah banyak dikenal dan digunakan untuk klasifikasi data yang memiliki atribut-atribut numerik dan kategorial. Hasil dari proses klasifikasi yang berupa aturan-aturan dapat digunakan untuk memprediksi nilai atribut bertipe diskret dari *record* yang baru (Angga Raditya, 2010).

Analisa Data

Data dalam pohon keputusan biasanya dinyatakan dalam bentuk tabel dengan *atribut dan record*. *Atribut* menyatakan suatu parameter yang dibuat sebagai kriteria dalam pembentukan pohon. Salah satu atribut merupakan *atribut* yang menyatakan data solusi per *item* data yang disebut target atribut. *Atribut* memiliki nilai-nilai yang dinamakan *instance*.

Pada penulisan penelitian ini, proses klasifikasi dilakukan untuk mengetahui kriteria nasabah non muslim menabung di bank syariah khususnya di

PT. Bank BRI Syariah Kantor Cabang Medan diproses dengan membuat pohon keputusan (*decision tree*) untuk menghasilkan suatu *output*. Data-data tersebut mempunyai beberapa kriteria yang merupakan syarat dalam pengklasifikasian kriteria nasabah non muslim menabung di bank syariah dan merupakan data input. Kemudian data yang telah diperoleh dianalisa dan dikelompokkan menjadi beberapa kelompok data yang dapat diproses dengan merancang pohon keputusan.

Penerapan *Data Mining*

Banyak penerapan yang dapat dilakukan oleh *data mining*, apalagi ditunjang dengan kekayaan dan keanekaragaman berbagi bidang ilmu seperti *artificial intelligence*, *database*, statistik, pemodelan matematika, pengolahan citra, dan lain-lain, membuat penerapan data mining menjadi makin luas, salah satunya dalam bidang perbankan.

Pemilihan Variabel

Dari data-data yang telah diperoleh, maka akan ditentukan satu variabel yang menjadi variabel keputusan yaitu nasabah non muslim yang menabung di bank syariah dan yang tidak menabung. Yang dimaksud dengan nasabah tidak menabung adalah nasabah tersebut menggunakan produk bank syariah bukan sebagai nasabah penabung tetapi menggunakan produk-produk lain seperti pembiayaan, jasa, dan lain-lain. Data yang menjadi variabel penentu dalam pembentukan pohon keputusan pada penelitian ini adalah jenis kelamin, umur, agama, pendidikan terakhir dan pekerjaan.

Pra Proses

Berdasarkan variabel-variabel yang sudah terpilih, maka diperoleh format data pada tabel 4.1.

Tabel 4.1 Data-Data Nasabah Non Muslim PT. Bank BRI Syariah Kantor Cabang Medan

No	Jenis klmn	Umur	Agama	Pendidikan terakhir	Pekerjaan	Jenis produk
1	Laki-laki	22 tahun	Kristen	Sma	Mahasiswa	Tab. Pendidikan
2	Laki-laki	19	Budha	Sma	Mahasiswa	Tab.

		tahun				Pendidikan
3	Laki-laki	25 tahun	Kristen	D3 ekonomi	Wirausaha	Tabungan
4	Perempuan	30 tahun	Kristen	S1 manaj	Ibu rumah tangga	Tabungan
5	Perempuan	35 tahun	Hindu	Sma	Ibu rumah tangga	Tabungan
6	Laki-laki	18 tahun	Budha	Sma	Pelajar	Tab. Pendidikan
7	Laki-laki	25 tahun	Budha	D3 komp.	Wirausaha	Giro
8	Perempuan	26 tahun	Kristen	S1 hukum	Pns	Tabungan
9	Laki-laki	27 tahun	Kristen	S1 ekonomi	Pns	Tabungan
10	Perempuan	40 tahun	Budha	Sma	Ibu rumah tangga	Tabungan
11	Perempuan	35 tahun	Budha	Sma	Ibu rumah tangga	Tabungan
12	Laki-laki	33 tahun	Kristen	S1 ekonomi	Wirausaha	Jasa (lc)
13	Perempuan	31 tahun	Kristen	S1 komp.	Wirausaha	Tabungan
14	Perempuan	22 tahun	Budha	Sma	Mahasiswa	Tab. Pendidikan
15	Laki-laki	45 tahun	Budha	Sma	Wirausaha	Pembiayaan
16	Perempuan	30 tahun	Budha	Sma	Ibu rumah tangga	Tabungan
17	Perempuan	50 tahun	Kristen	D3 akuntns	Pns	Tabungan
18	Laki-laki	30 tahun	Hindu	S1 ekonomi	Wirausaha	Giro

19	Laki-laki	20 tahun	Kristen	Sma	Mahasiswa	Tab. Pendidikan
20	Perempuan	33 tahun	Hindu	S1 akuntansi	Wirausaha	Giro

Setelah diketahui data dalam format di atas, kemudian dilakukan pra-proses dengan mengelompokkan data-data tersebut ke dalam kelas, kemudian dilakukan pra proses dengan cara mengelompokkan variabel-variabel penentu sebagai berikut:

1. Mengelompokkan nilai jenis kelamin, pengelompokkan jenis kelamin ini dilakukan pengujian berdasarkan jenis kelamin nasabah sehingga nilai tersebut dikelompokkan seperti pada tabel 4.2
2. Mengelompokkan nilai umur, pengelompokkan umur ini dilakukan pengujian berdasarkan *Transformasi Algoritma MD Heuristik* dengan tahapam sebagai berikut:

a. Bentuk *Set Of Cuts*

$$a(U) = \{18, 19, 20, 22, 25, 26, 27, 30, 31, 33, 34, 35, 40, 41, 45, 50\}$$

$$VB(A) = \{ Pa1, Pa2, Pa3, Pa4, Pa5, Pa6, Pa7, Pa8, Pa9, Pa10, Pa11, Pa12, Pa13, Pa14, \}$$

Dimana nilai variable *set of cuts* tertera pada table 4.3 berikut :

Tabel 4.3 Variabel *Set Of Cuts*

Variabel	Nilai
<i>Pa1</i>	18 – 19
<i>Pa2</i>	19 – 20
<i>Pa3</i>	20 – 22
<i>Pa4</i>	22 – 25
<i>Pa5</i>	25 – 26
<i>Pa6</i>	26 – 27
<i>Pa7</i>	27 – 30
<i>Pa8</i>	30 – 31
<i>Pa9</i>	31 – 33
<i>Pa10</i>	33 – 35

<i>Pa 11</i>	35 – 40
<i>Pa 12</i>	40 – 41
<i>Pa 13</i>	41 – 45
<i>Pa 14</i>	45 – 50

Hasil Pengujian

Pengujian terhadap hasil analisa data, sangat penting dilakukan untuk menentukan dan memastikan apakah hasil dari sebuah analisa tersebut benar atau tidak.

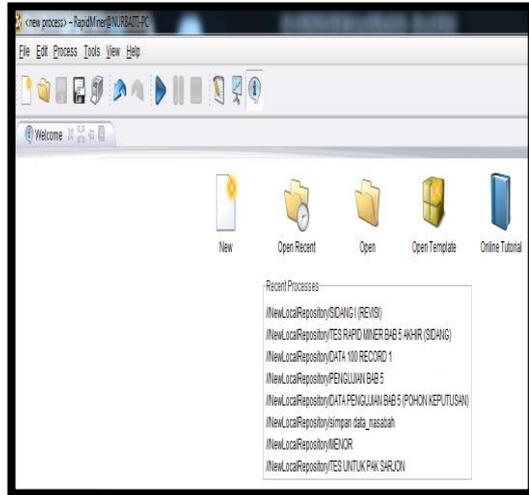
Proses pengujian dilakukan untuk mengetahui kebenaran dari hasil pengolahan data yang dikerjakan secara manual, untuk proses pengujian tersebut kita dapat menggunakan salah satu *software* aplikasi seperti *Rapidminer*. Adapun langkah-langkah yang dilakukan dalam pengujian menggunakan *software Rapidminer* adalah sebagai berikut:

1. *Dataset* yang diuji terdiri dari 100 *record* data dan berisi variabel-variabel atribut kondisi yang terdiri dari variabel jenis kelamin, umur, agama, pendidikan terakhir dan pekerjaan, sedangkan atribut keputusannya adalah menabung dan tidak menabung. Data nasabah tersebut disimpan dalam aplikasi *microsoft excel* dengan nama *file* Data Pengujian Bab 5. *xls* seperti terlihat pada Tabel 5.1.

No	Jenis klmn	Umur	Agama	Pendidikan terakhir	Pekerjaan	Keputusan
1	Lk	1	Kristen	Sma	Wiraha	Ya
2	Lk	1	Budha	Sma	Pelajar	Ya
3	Lk	2	Kristen	Diploma	Wiraha	Ya
4	Pr	4	Kristen	Sarjana	Pelajar	Ya
5	Pr	7	Hindu	Sma	Irt	Ya
6	Lk	1	Budha	Sma	Pelajar	Ya
7	Lk	2	Budha	Diploma	Wiraha	Tidak
8	Pr	2	Kristen	Sarjana	Pns	Ya

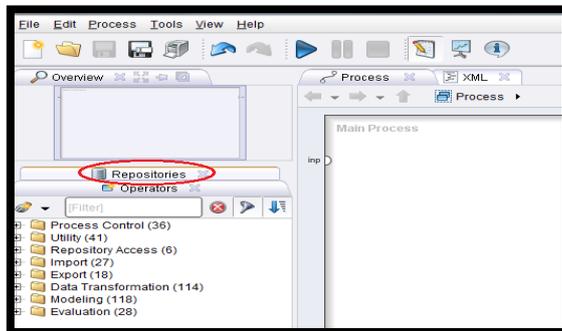
9	Lk	2	Kristen	Sarjana	Pns	Ya
10	Pr	7	Budha	Sma	Irt	Ya

2. Buka software *Rapidminer*



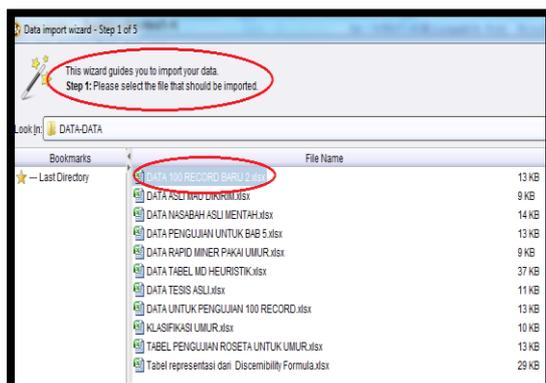
Gambar 5.1 Tampilan Menu Awal *Rapidminer*

Untuk memulai klik *Menu File*, Pilih *New* dan pilih tab *Repositories* sehingga muncul tampilan seperti:



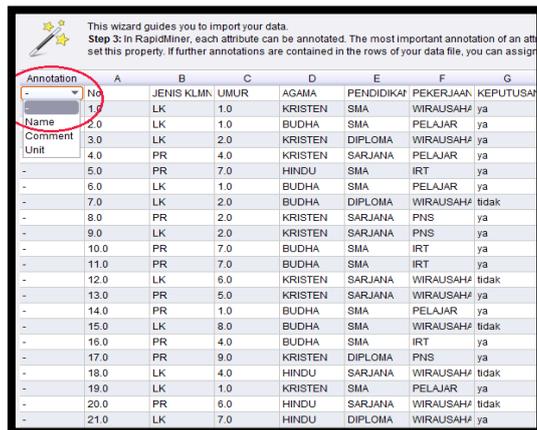
Gambar 5.2 Tampilan *Rapidminer Repositories*

3. Pilih Menu *File*, pilih *Import Data* dan pilih *Import Excel Sheet*, sehingga akan muncul tampilan seperti berikut :



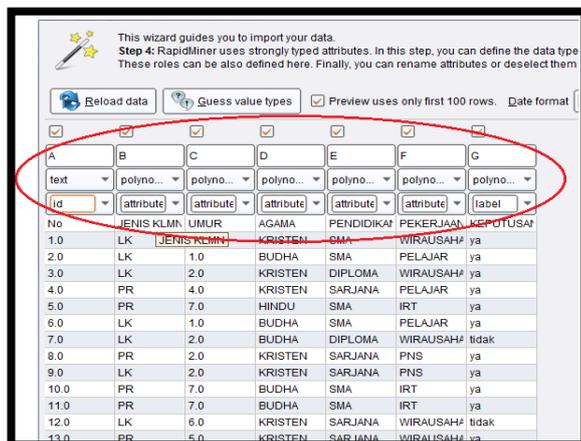
Gambar 5.3 Tampilan *Import Excel Sheet*

4. Arahkan ke *folder* tempat *file* Data Pengujian Bab 5.xls disimpan, kemudian klik pada file Data Pengujian Bab 5.xls kemudian klik tombol *next* dan klik *Next* sekali lagi.
5. Klik pada baris pertama di kolom *Annotation*, pilih *name* dan klik *next* seperti tampilan Gambar 5.4



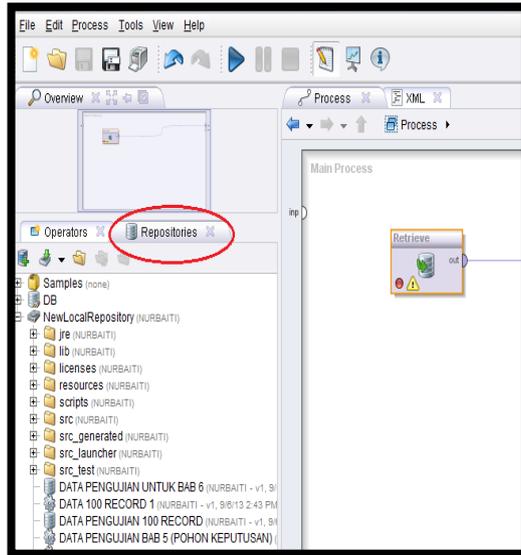
Gambar 5.4 Pilihan *Annotation*

6. Selanjutnya ubah properties atribut nomor menjadi *id* dan atribut keputusan menjadi *label*, dan sesuaikan tipe tiap *field* atribut lalu klik *Next* seperti tampilan berikut :



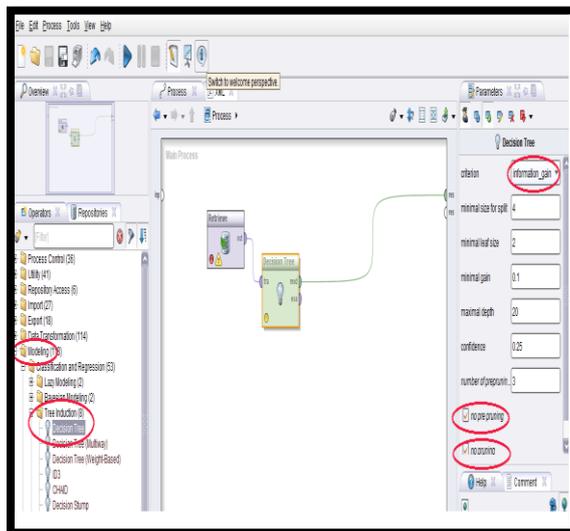
Gambar 5.5 Pilihan *Properties Atribut* dan *Field*

7. Ketikkan Data Pengujian Bab 5 di *textbox* *Name*, kemudian klik tombol *Finish* untuk menyimpan *dataset* yang akan kita pakai.
8. Untuk menghasilkan pohon keputusan dengan algoritma C4.5, pilih *dataset* yang kita sudah simpan sebelumnya di *Tab Repositories*, kemudian *drag dataset* data uji ke *Tab Process*, seperti tampilan gambar 5.6



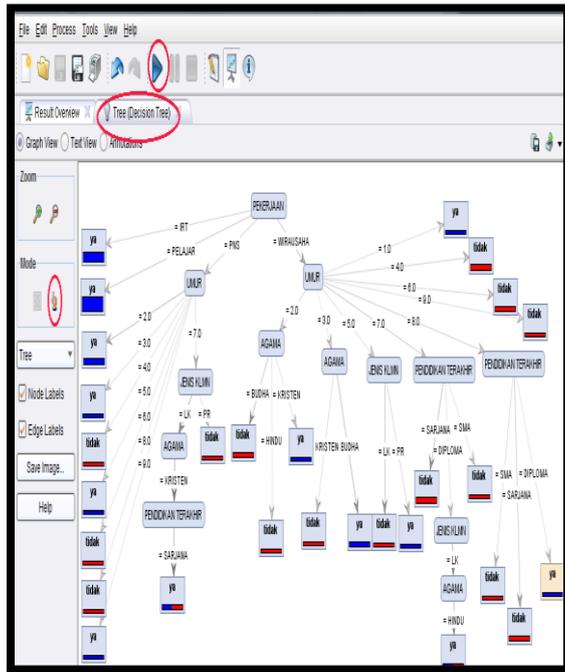
Gambar 5.6 Tampilan *Repositories*

9. Selanjutnya tambahkan algoritma *decision tree* dari *Tab Operators*, pilih *Modelling (118)*, pilih *Classification And Regression (53)*, pilih *Tree Induction (8)*, kemudian *drag Decision Tree* ke *Tab Process*, pada *Tab Parameters* pilih *information_gain* pada *criterion*, dan cek pada pilihan *no pre pruning* dan *no pruning*, seperti gambar 5.7.



Gambar 5.7 Tampilan *Process Decision Tree*

10. Pilih menu *Process* dan pilih *Run*, pilih *No* dan pilih *Yes*, secara otomatis Rapidminer akan memproses untuk menghasilkan pohon keputusan seperti gambar 5.8.



Gambar 5.8 Tampilan *Rapidminer* Pohon

Kesimpulan

Dari uraian pada bab-bab yang sudah dibahas sebelumnya dapat ditarik kesimpulan:

1. Metode pohon keputusan (*decision tree*) yang diproses dengan *software Rapidminer* lebih efektif dan fleksibel jika digunakan pada proses pengklasifikasian nasabah non muslim di bank syariah.
2. Pemilihan variabel (atribut kondisi dan attribut keputusan) yang akan digunakan dalam menentukan sebuah klasifikasi juga sangat mempengaruhi *rule* atau *knowledge* yang dihasilkan.
3. Sistem yang dibangun dapat membantu dalam mengklasifikasikan atau kriteria nasabah non muslim yang menggunakan produk bank syariah yang berguna bagi pihak manajemen atau marketing bank syariah dalam pengembangan jumlah atau volume nasabah bank.
4. Sistem pengklasifikasian nasabah non muslim menggunakan algoritma *C4.5* dapat digunakan dalam pengambilan keputusan untuk mencari solusi penetapan bentuk promosi atau pelayanan yang tepat guna untuk pengembangan jumlah nasabah.
5. Algoritma *C4.5* dianggap sebagai algoritma yang sangat membantu dalam melakukan klasifikasi data karena karakteristik data yang diklasifikasi

dapat diperoleh dengan jelas, baik dalam bentuk struktur pohon keputusan (*decision tree*) maupun dalam aturan rule *If – Then* sehingga memudahkan pengguna dalam melakukan penggalian informasi terhadap data yang bersangkutan.

Saran

Pengembangan sistem ini adalah dari sistem manual atau berdasarkan basis pengetahuan ke sistem komputerisasi yang dibangun atas dasar pertimbangan permasalahan-permasalahan yang timbul dari sistem yang lama. Berikut ini adalah saran yang dapat diperhatikan untuk masa yang akan datang.

1. Pada penelitian ini, penulis mencoba salah satu teknik yang digunakan dalam mengklasifikasikan nasabah yaitu dengan teori *decision tree*. Untuk mendapatkan hasil yang baik dapat digunakan beberapa atau penggabungan beberapa teknik klasifikasi. Oleh sebab itu perlunya dilakukan penelitian lanjut untuk membandingkan hasil klasifikasi yang lain dan menentukan teori mana yang menghasilkan klasifikasi yang baik.
2. Untuk mendapatkan hasil klasifikasi yang baik diperlukan sumber data yang lengkap. Oleh sebab itu diharapkan pada proses penyimpanan data nasabah di sebuah bank harus dilakukan secara elektronik dengan demikian akan memudahkan dalam pencarian data yang dibutuhkan untuk melakukan proses penemuan *knowledge* seperti untuk mengetahui jenis-jenis kriteria nasabah di sebuah perbankan.
3. Walaupun aplikasi ini sangat mudah untuk di jalankan, namun perlu diberitahukan cara-cara pengoperasiannya, sehingga pihak manajemen perbankan khususnya bank syariah bisa langsung dapat menggali data-data nasabahnya dengan menggunakan *software* ini.

Daftar Pustaka

- Afaq Alam Khan¹, Sanjay Jamwal², M.M Sepehri³. 2010. *Applying Data Mining to Customer Churn Prediction in an Internet Service Provider*.
- Ahmad Basuki¹, Iwan Syarif². 2003. *Ebook*. Politeknik Elektronika Negeri Surabaya.

- Amritpal Kaur¹, Sandeep Singh². 2013. *Classification and Selection of Best Saving Service for Potential Investors using Decision Tree – Data Mining Algorithms*.
- Angga Ginanjar Mabur¹, Riani Lubis². 2012. *Penerapan Data Mining Untuk Memprediksi Kriteria Nasabah Kredit*.
- Angga Raditya. 2010. *Implementasi Data Mining Classification untuk Mencari Pola Prediksi Hujan dengan Menggunakan Algoritma C4.5*.
- Brijesh Kumar Bhardwaj¹, Saurabh Pal². 2011. *Data Mining : A Prediction for Performance Improvement Using Classification*.
- Daniel Larose. 2005. *Discovering Knowledge in Data (An Introduction to Data Mining)*. United States of America : Wiley Interscience
- Dedy Rahman Wijaya. 2005. *Data Mining Introduction*.
- Fajar Astuti Hermawati. 2013. *Data Mining*. Yogyakarta : Andi
- Hamidah Jantan¹, Abdul Razak², Zulaiha Ali Othman³. 2010. *Human Talent Prediction In HRM Using C4.5 Classification Algorithm*.
- Jiawei Han¹, Micheline Kamber², Jian Pei³. 2006. *Data Mining Concepts and Techniques*.
- K. Rajesh¹, Sheila Anand². 2012. *Analysis of SEER Dataset for Breast Cancer Diagnosis using C4.5 Classification Algorithm*.
- Khalida Binti Oseman¹, Sunartati², Norazrina³, Faizin⁴. 2010. *Data Mining in Churn Analysis Model for Telecommunication Industry*.
- Kusrini dan Luthfi Taufiq Emha. 2009. *Algoritma Data Mining*. Yogyakarta : Andi
- Max Bramer. 2007. *Principles of Data Mining*.
- Michael J. A. Berry¹, Berry Gordon S. Linoff². 2004. *Data Mining Techniques*.
- Sunjana. 2010. *Aplikasi Mining Data Mahasiswa dengan Metode Klasifikasi Decision Tree*.
- Yudho Giri Sucahyo. 2003. *Data Mining Menggali Informasi yang Terpendam*.