

IMPLEMENTASI METODE DECISION TREE DAN ALGORITMA C4.5 UNTUK KLASIFIKASI DATA NASABAH BANK

Rusito¹, Meidy Taufany Firmansyah²

^{1, 2} Program Studi S1 Sistem Komputer STEKOM Semarang

rusito@stekom.ac.id

Abstrak

Salah satu keberhasilan perbankan adalah mengklasifikasi nasabahnya. Kegiatan klasifikasi erat kaitan dengan dokumen profing. *Proofing* adalah salah satu kegiatan pencatatan dokumen jaminan yang nantinya data-data tersebut akan di klasifikasikan sesuai dengan ketentuan yang telah ditetapkan oleh pihak manajemen. Kegiatan ini berfungsi sebagai pengumpulan data-data nasabah sebagai kreditur. Metode yang digunakan dalam sistem ini menggunakan metode *decision tree* algoritma C.45 dan merupakan salah satu metode yang ada pada klasifikasi dalam *data mining*. Ada tiga tahap dalam pembuatan aplikasi ini, yaitu tahapan pengolahan data, tahap *decision tree*, dan tahap interpretasi. Pada tahap pengolahan data meliputi data *selection*, tahap *decision tree* ialah tahap pembuatan tree, sedangkan tahap interpretasi adalah proses pencarian jenis dokumen berdasarkan *decision tree* yang telah didapat. penulis menggunakan teori *Research and Development (R&D)* atau *research-based development*, yaitu pengembangan berbasis penelitian yaitu proses yang digunakan untuk mengembangkan dan memvalidasi produk-produk pemrograman. Hasil dari penelitian ini adalah sebuah produk sistem yang telah di validasi pakar materi (*Reviewer*) dengan uji validasi nilai 3,7 yang berarti valid. Karena terletak pada kriteria kevali dan 2,51 – 3,25. Dan pakar materi (*Reviewer*) menyimpulkan bahwa sistem layak dipakai.

Kata kunci : *Decision tree*, Algoritma C4.5, Klasifikasi, Nasabah

1. PENDAHULUAN

Dunia telah berubah dalam era digital, perubahan besar dapat dilihat pada segala bidang, terutama penyampaian informasi yang lebih sistematis dan informatif. Demikian juga pada bidang lain seperti perbankan dengan eBanking, pemerintahan dengan eGoverment dan lain sebagainya. Perubahan ke era digital mengubah bentuk data dan media penyimpanan selama ini, data-data yang tersimpan tidak lagi dalam bentuk kertas.

Pencatatan dokumen pada bagian penyimpanan dokumen jaminan atau disebut *Loan Doc and Safe Keeping (LDSK)*, saat ini masih kebingungan dalam hal menentukan kelayakan seberapa besar kredit yang telah dikucurkan kepada nasabah apakah telah sesuai dengan jaminan yang di telah agunkan oleh nasabah. Sehingga diperlukan kejelian pada saat proses pembuatan laporan, karena jika tidak teliti maka akan terjadi kerancuan saat menentukan kelas nasabah tersebut. Seluruh dokumen juga masih terkumpul menjadi satu, belum terkelaskan sesuai dengan account pembukaan rekening, sehingga perlu memilih-milih lagi dokumen antar bagian. Pencatatan dukumen tersebut erat kaitanya dengan klasifikasi.

Salah satu keberhasilan perbankan adalah mengklasifikasi nasabahnya. Kegiatan klasifikasi erat kaitan dengan dokumen profing. *Proofing* adalah salah satu kegiatan pencatatan dokumen jaminan yang nantinya data-data tersebut akan di klasifikasikan sesuai dengan ketentuan yang telah ditetapkan oleh pihak manajemen. Kegiatan ini berfungsi sebagai pengumpulan data-data nasabah sebagai kreditur.

Penerapan Data Mining di Bidang Perbankan, Bank memiliki masalah memprediksi kelayakan kredit dari nasabah baru berdasarkan data historis dari nasabah masa lalu. Bagian ini menjelaskan bagaimana Data Mining dapat diterapkan untuk masalah ini. Sebuah bank memiliki data tentang nasabah kepada

siapa itu memberi kredit di masa lalu. Data nasabah berisi data pribadi, data yang menjelaskan status keuangan dan perilaku keuangan sebelum dan pada saat itu nasabah diberi kredit [1].

Pemanfaatan teknik data mining menggunakan metode algoritma C4.5 dalam mengklasifikasikan data nasabah diharapkan dapat memberikan informasi-informasi yang sebelumnya tersembunyi didalam gudang data sehingga menjadi informasi yang berharga. [1]

Penelitian ini merancang sebuah Sistem Informasi Klasifikasi Data Nasabah menggunakan metode Data Mining dan Algoritma C4.5 berbasis web, sehingga bisa diakses secara besamaan di bagian LDSK atau antar bagian dan bertujuan untuk membantu menyelesaikan masalah kelayakan dalam pemberian kredit. Dokumen agunan nantinya akan dibagi menjadi tiga kelas. Dokumen kelas pertama adalah nasabah kelas atas (High End), Dokumen kelas kedua adalah nasabah kelas menengah (Middle), dan Dokumen kelas adalah nasabah kecil (Small).

2. KAJIAN LITERATUR

2.1. Klasifikasi

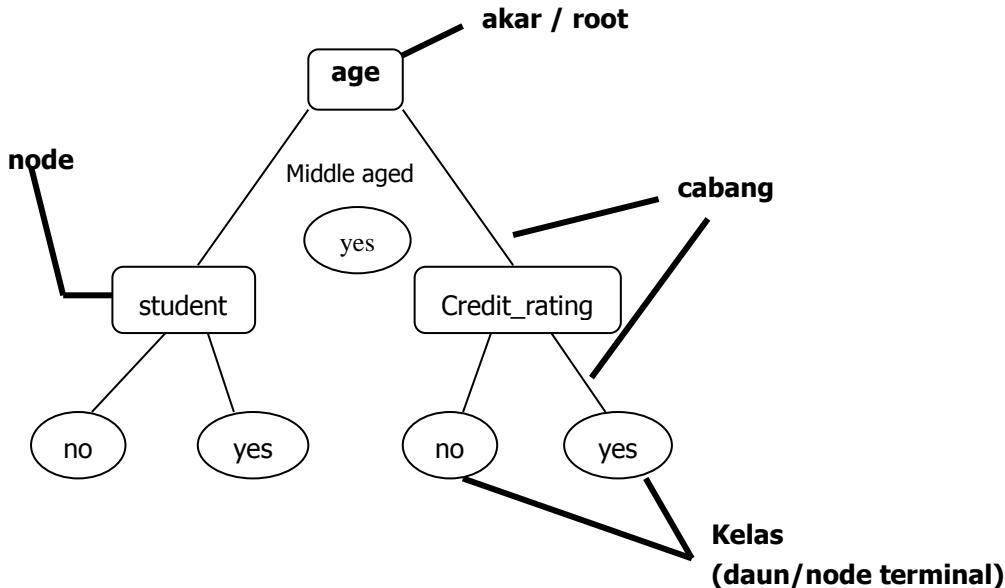
Klasifikasi dokumen adalah pemberian kategori yang telah didefinisikan ke dokumen. Mengklasifikasikan dokumen merupakan salah satu cara untuk mengorganisasikan dokumen. Dokumen-dokumen yang memiliki isi yang sama akan dikelompokkan ke dalam kategori yang sama [6]. Dengan demikian orang-orang yang melakukan pencarian informasi akan tertolong dalam hal navigasi dan dengan mudah kategori yang tidak relevan dengan informasi dapat segera dilewatkan [7].

Pada awalnya, kegiatan klasifikasi dokumen dilakukan secara manual oleh seorang ahli. Untuk melakukan klasifikasi, ahli harus menyediakan waktu untuk membaca isi setiap dokumen baru dan membandingkan isinya dengan skema kategori yang tersedia. Dengan demikian klasifikasi ini merupakan kegiatan yang menghabiskan waktu serta biaya. Selain alasan yang telah yang telah disebutkan sebelumnya ditambah dengan peningkatan jumlah dokumen telah mendorong berkembangnya metode pengklasifikasian secara otomatis yang dapat melakukan klasifikasi dengan cara belajar dari sekumpulan contoh dokumen yang telah diklasifikasikan sebelumnya. Dengan metode tersebut, keuntungan yang diperoleh adalah penghematan tenaga kerja dan efektivitas yang baik.

2.2. DECISION TREE

Metode ini adalah sebuah struktur pohon, dimana setiap node pohon merepresentasi-kan atribut yang telah diuji, setiap cabang merupakan suatu pembagian hasil uji, dan node daun (leaf) merepresentasikan kelompok kelas tertentu. Level node teratas dari sebuah Decision Tree adalah node akar (root) yang biasanya berupa atribut yang paling memiliki pengaruh terbesar pada suatu kelas tertentu. Pada umumnya Decision Tree melakukan strategi pencarian secara top-down untuk solusinya. Pada proses mengklasifikasi data yang tidak diketahui, nilai atribut akan diuji dengan cara melacak jalur dari node akar (root) sampai node akhir (daun) dan kemudian akan diprediksi kelas yang dimiliki oleh suatu data baru tertentu. Secara singkat bahwa Decision Tree merupakan salah satu metode klasifikasi pada Text Mining.

Seperti ditunjukkan dalam Gambar 1, decision tree tergantung pada aturan if-then, tetapi tidak membutuhkan parameter dan metric. Strukturnya yang sederhana dan dapat ditafsirkan memungkinkan decision tree untuk memecahkan masalah atribut multi-type. Decision tree juga dapat mengelola nilai-nilai yang hilang atau data noise.[2]



Gambar 1. Struktur Pohon Keputusan (Han, 2001)

Dari pohon keputusan diatas maka didapatkan aturan klasifikasi sebagai berikut [2]:

- R1: **IF** age = youth **AND** student = no **THEN** buys computer = no
- R2: **IF** age = youth **AND** student = yes **THEN** buys computer = yes
- R3: **IF** age = middle aged **THEN** buys computer = yes
- R4: **IF** age = senior **AND** credit rating = excellent **THEN** buys computer = yes
- R5: **IF** age = senior **AND** credit rating = fair **THEN** buys computer = no

2.3. ALGORITMA C4.5

Algoritma C4.5 merupakan algoritma yang digunakan untuk membentuk pohon keputusan. Pohon keputusan merupakan metode klasifikasi dan prediksi yang sangat kuat dan terkenal. Metode pohon keputusan mengubah fakta yang sangat besar menjadi pohon keputusan yang merepresentasikan aturan. Aturan dapat dengan mudah dipahami dengan bahasa alami. Dan mereka juga dapat diekspresikan dalam bentuk bahasa basis data seperti Structured Query Language untuk mencari record pada kategori tertentu. Pohon Keputusan juga berguna untuk mengeksplorasi data, menemukan hubungan tersembunyi antara sejumlah calon variabel input dengan sebuah variabel target. [3]

Ada beberapa tahapan dalam membuat sebuah pohon keputusan dalam algoritma C4.5, yaitu :

1. Mempersiapkan data nasabah. Data nasabah diambil dari record yang ada sebelumnya atau disebut data masa lalu dan sudah dikelompokkan dalam kelas-kelas tertentu.
2. Menghitung akar dari pohon. Akar akan diambil dari atribut yang akan terpilih, 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.
 - Tahapan Algoritma Decision Tree C4.5 :
 - Menyiapkan data training
 - Menentukan akar dari pohon
 - Hitung nilai Gain :

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

S : Himpunan Kasus

n : Jumlah partisi S

pi : Proporsi dari Si terhadap S

- Menghitung informasi Gain

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

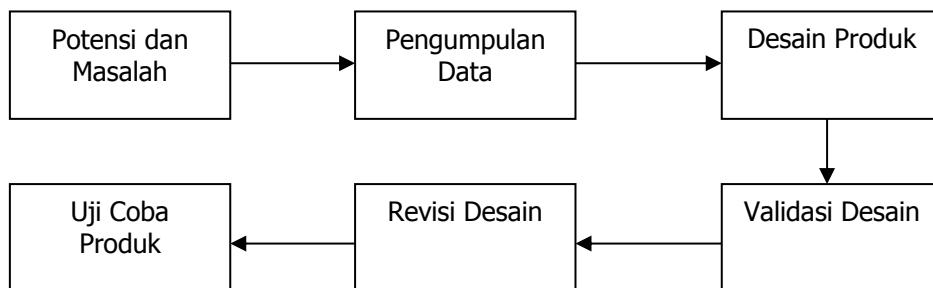
- S : Himpunan kasus
 A : Atribut
 n : Jumlah partisi atribut A
 $|S_i|$: Jumlah kasus pada partisi ke i
 $|S|$: Jumlah kasus dalam S
-

3. METODE PENELITIAN

Dalam hal prosedur penelitian pengembangan Borg & Gall (1987) mengungkapkan bahwa R&D tersusun dalam beberapa langkah penelitian sebagai berikut: [5] penelitian dan pengumpulan infomasi (research and information collecting); perencanaan (planning), pengembangan produk pendahuluan (develop premilitary form of product); uji coba pendahuluan (premilitary field testing); perbaikan produk utama (main product revision); uji coba utama (main field testing); perbaikan produk operasional (operational product revision); uji coba lapangan (operational field testing); perbaikan produk akhir (final product revision); diseminasi dan pendistribusian (dissemination and distribution). [4]

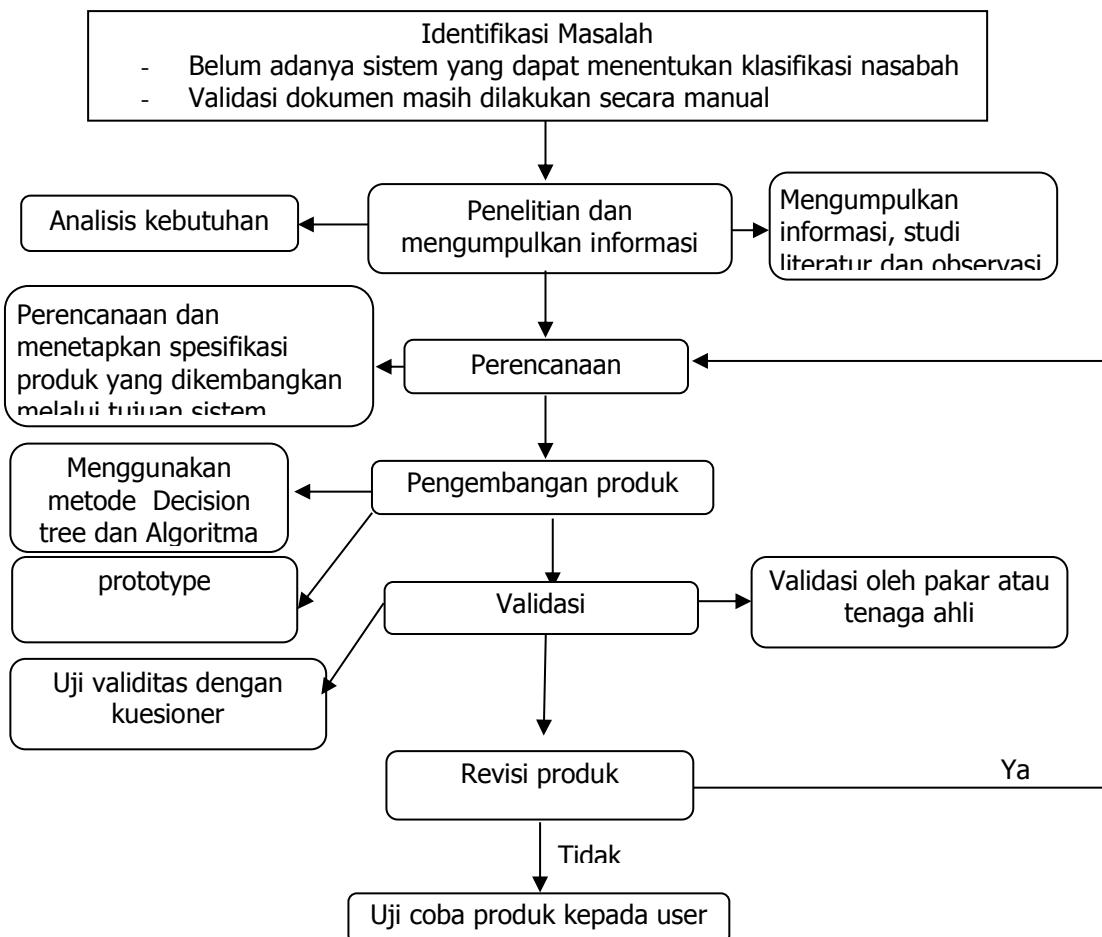
Penelitian ini mengacu pada langkah-langkah yang dilakukan oleh borg & gall yang kemudian dimodifikasi menjadi 6 langkah sebagai berikut:

- Analisis potensi dan masalah
- Pengumpulan data dan informasi
- Desain produk
- Validasi design
- Perbaikan desain
- Uji coba produk



Gambar 2. Langkah Tahapan Pengembangan Research and Development (Sugiyono, 2012)

Kerangka Berfikir



Gambar 3. Kerangka Berfikir

4. HASIL PENELITIAN

4.1. Perhitungan Metode Decision Tree dan Algoritma C4.5

Dari data table nasabah, maka dilakukan perhitungan menggunakan metode Decision Tree dan Algoritma C4.5, agar di dapat jenis klas nasabah. Berdasarkan data tabel nasabah, membuat tabel percabangan. Adapaun calon percabangan yang mungkin terjadi dari tabel pertama adalah :

Tabel 1. Tabel Calon Percabangan

PRESTASI	JAMINAN	PEKERJAAN
BAIK	SERTIFIKAT HM	PNS
BERMASALAH	BPKB	SWASTA
	LOGAM MULIA	WIRASWASTA

Tabel 2. Tabel Sampel Nasabah

No	No rek.	Nama	Alamat	Prestasi	Jaminan	Pekerjaan	Klas
1	4530100061814	Adityo	Jl brotojoyo timur 8 smg	Baik	Sertifikat HM	PNS	Atas
2	4530100061814	Christanto	Jl.dr.cipto no.125 semarang	Bermasalah	Sertifikat HM	PNS	Atas
3	0161100003340	Agung wijanarko	Jl amarilis sungai ix/58 perum graha perdana indah	Baik	BPKB	Swasta	Sedang
4	2290100536007	Ita kristiana	Jl dr sutomo 47 smg	Baik	Logam mulia	Swasta	Mikro
5	453110000016	Kamari	Jl.gemah raya ii/60 semarang	Baik	BPKB	Wiraswasta	Atas
6	2290100342000	Lianawati soegiharto	Jl.brigjend katamso no.15 purwodadi	Bermasalah	BPKB	Wiraswasta	Mikro
7	016120000287	Suharnanto	Kompl.pangkalan truk blok b.22 smg	Baik	Sertifikat HM	Swasta	Sedang
8	4530100055813	Victor usman	Jl. Srondol bumi indah blok i.1 semarang	Baik	Sertifikat Hm	Swasta	Atas
9	229090000441	Zainul ibad	Jl. Tirto 112 agung, semarang	Bermasalah	Logam Mulia	Swasta	Mikro

Berikut membuat tabel untuk menghitung frekwensi yang muncul dari setiap hasil produk berikut dengan proporsinya serta entropinya $H(t)$

Tabel 3. Tabel Frekwensi

KELAYAKAN	FREK	Pi	Log2 Pi	-Pi.log2Pi
ATAS	4	0,444444444	-0,416735694	0,185215864
SEDANG	2	0,222222222	-0,772942881	0,171765085
MIKRO	3	0,333333333	-0,564575034	0,188191678

Tabel 4. Tabel Perhitungan Gain

CABANG	Jml Kasus (S)	KELAS ATAS (S1)	KELAS SEDANG (S2)	KELAS MIKRO (S3)	Entropy (Si)	GAIN
KREDIT BERMASALAH	BAIK	6	3	2	1	-0,73369
	BERMASALAH	3	1	0	2	
JAMINAN	SERTIFIKAT HM	4	3	1	0	-0,34372
	BPKB	3	1	1	1	
	LOGAM MULIA	2	0	0	2	
PEKERJAAN	PNS	2	2	0	0	-0,52257
	SWASTA	5	1	2	2	
	WIRASWASTA	2	1	0	1	

Penghitungan Entropy Kelayakan

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

dengan :

S : Himpunan Kasus

n : Jumlah partisi S

pi : Proporsi dari Si terhadap S

$$\text{Entropy } (\text{Total}) = \left(-\frac{4}{9}x \log_2 \left(\frac{4}{9}\right)\right) + \left(-\frac{2}{9}x \log_2 \left(\frac{2}{9}\right)\right) + \left(-\frac{3}{9}x \log_2 \left(\frac{3}{9}\right)\right)$$

$$\text{Entropy } (\text{Total}) = (-0,4444 \times -0,416735694) + (-0,2222 \times -0,772942881) + (-0,3333 \times -0,564575034)$$

$$\text{Entropy } (\text{Total}) = 0,185215864 + 0,171765084 + 0,188191678$$

$$\text{Entropy } (\text{Total}) = 0,545172627$$

Rumus penghitungan informasi Gain

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

dengan :

S : Himpunan kasus

A : Atribut

n : Jumlah partisi atribut A

|Si| : Jumlah kasus pada partisi ke i

|S| : Jumlah kasus dalam S

Penghitungan informasi Gain Prestasi

$$\begin{aligned}
 S_{baik} [3, 2, 1] &= (-\frac{3}{6}x \log_2(\frac{3}{6})) + (-\frac{2}{6}x \log_2(\frac{2}{6})) + (-\frac{1}{6}x \log_2(\frac{1}{6})) \\
 &= 0,5 + 0,5283 + 0,4308 \\
 &= 1,45915 \\
 S_{bermasalah} [1, 0, 2] &= (-\frac{1}{3}x \log_2(\frac{1}{3})) + (-\frac{0}{3}x \log_2(\frac{0}{3})) + (-\frac{2}{3}x \log_2(\frac{2}{3})) \\
 &= 0,5283 + 0 + 0,389975 \\
 &= \mathbf{0,9183}
 \end{aligned}$$

$$\begin{aligned}
 Gain (S, Prestasi) &= Entropy S - (6/9) S_{baik} - (3/9) S_{bermasalah} \\
 &= 0,545172627 - (6/9) \times 1,45915 - (3/9) \times 0,9183 \\
 &= 0,545172627 - 0,9727666 - 0,3061 \\
 &= \mathbf{-0,73369}
 \end{aligned}$$

Penghitungan informasi Gain Jaminan

$$\begin{aligned}
 SSERTIFIKAT HM [3, 1, 0] &= (-\frac{3}{4}x \log_2(\frac{3}{4})) + (-\frac{1}{4}x \log_2(\frac{1}{4})) + (-\frac{0}{4}x \log_2(\frac{0}{4})) \\
 &= 0,311278124 + 0,5 + 0 \\
 &= 0,81128 \\
 SBPKB [1, 1, 1] &= (-\frac{1}{3}x \log_2(\frac{1}{3})) + (-\frac{1}{3}x \log_2(\frac{1}{3})) + (-\frac{1}{3}x \log_2(\frac{1}{3})) \\
 &= 0,528320834 + 0,528320834 + 0,528320834 \\
 &= 1,58496 \\
 SLOGAM MULIA [0, 0, 2] &= (-\frac{0}{2}x \log_2(\frac{0}{2})) + (-\frac{0}{2}x \log_2(\frac{0}{2})) + (-\frac{2}{2}x \log_2(\frac{2}{2})) \\
 &= 0 + 0 + 0 \\
 &= 0
 \end{aligned}$$

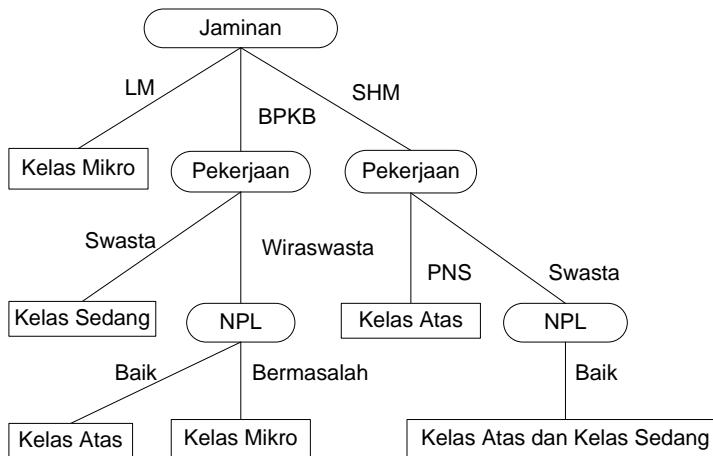
$$\begin{aligned}
 Gain (S, Jaminan) &= Entropy S - (4/9) SSERTIFIKAT HM - (3/9) SBPKB - (2/9) SLOGAM MULIA \\
 &= 0,545172627 - (4/9) \times 0,81128 - (3/9) \times 1,58496 - (2/9) \times 0 \\
 &= 0,545172627 - 0,360568055 - 0,528320833 - 0 \\
 &= \mathbf{-0,34372}
 \end{aligned}$$

Penghitungan informasi Gain Pekerjaan

$$\begin{aligned}
 SPNS [2, 0, 0] &= (-\frac{2}{2}x \log_2(\frac{2}{2})) + (-\frac{0}{2}x \log_2(\frac{0}{2})) + (-\frac{0}{2}x \log_2(\frac{0}{2})) \\
 &= 0 + 0 + 0 \\
 &= 0 \\
 SSwasta [1, 2, 2] &= (-\frac{1}{5}x \log_2(\frac{1}{5})) + (-\frac{2}{5}x \log_2(\frac{2}{5})) + (-\frac{2}{5}x \log_2(\frac{2}{5})) \\
 &= 0,464385619 + 0,528771238 + 0,528771238 \\
 &= 1,52193 \\
 SWiraswasta [1, 0, 1] &= (-\frac{1}{2}x \log_2(\frac{1}{2})) + (-\frac{0}{2}x \log_2(\frac{0}{2})) + (-\frac{1}{2}x \log_2(\frac{1}{2})) \\
 &= 0,5 + 0 + 0,5 \\
 &= 1
 \end{aligned}$$

$$\begin{aligned}
 Gain (S, Pekerjaan) &= Entropy S - (2/9) SPNS - (5/9) SSwasta - (2/9) SWiraswasta \\
 &= 0,545172627 - (2/9) \times 0 - (5/9) \times 1,52193 - (2/9) \times 1 \\
 &= 0,545172627 - 0 - 0,845515608 - 0,22222222 \\
 &= \mathbf{-0,52257}
 \end{aligned}$$

Berdasarkan hasil perhitungan information gain, tampak bahwa atribut jaminan akan menyediakan prediksi terbaik untuk target atribut kelas prioritas. Gambar berikut di bawah ini menjelaskan bahwa kriteria jaminan menjadi prioritas utama dalam menentukan hasil keputusan.

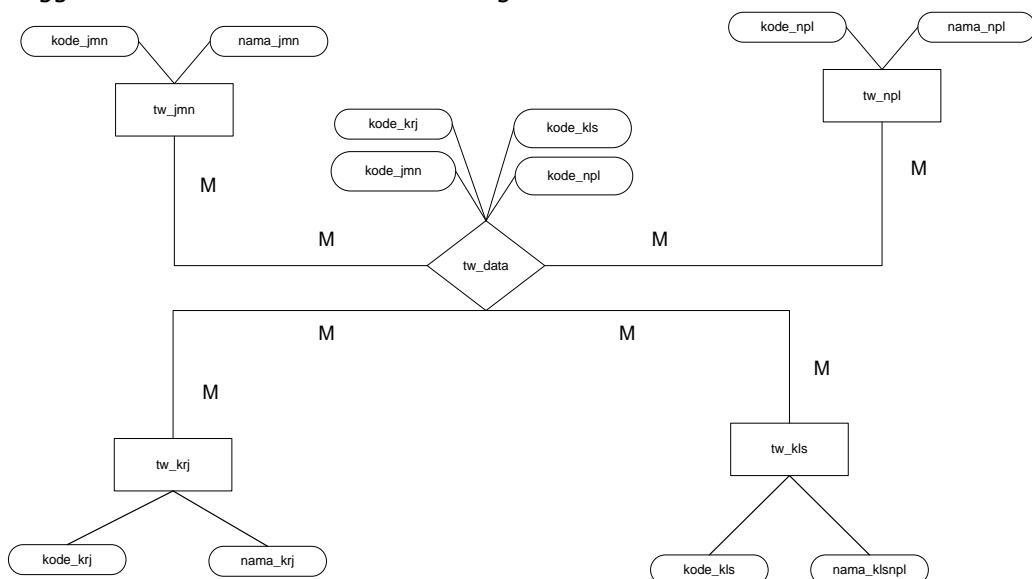
**Gambar 4. Tree Nasabah**

Dari gambar di atas terlihat dua simbol yaitu simbol kotak dan simbol elips, simbol kotak menandakan keputusan final artinya, karena nilai gain jaminan paling tinggi maka cabang jaminan menjadi cabang penentuan pertama yaitu sebagai berikut :

- 1) Jika jaminannya LOGAM MULIA maka akan mendapatkan tingkat kelas kelayakan mikro.
- 2) namun jika jaminan BPKB dan SERTIFIKAT HM tidak bisa menentukan tingkat kelas kelayakan, karena belum final keputusannya maka digunakan cabang kedua yang lebih tinggi nilai gainnya yaitu cabang pekerjaan dimana jaminan BPKB dengan pekerjaan Swasta mendapatkan tingkat kelas kelayakan sedang.
- 3) sedangkan jaminan BPKB dengan pekerjaan Wiraswasta belum mendapatkan keputusan final, maka digunakan cabang ketiga yaitu cabang Prestasi dimana jaminan BPKB dengan pekerjaan Wiraswasta dan Prestasi baik mendapatkan tingkat kelayakan kelas atas.
- 4) sedangkan jaminan BPKB dengan pekerjaan Wiraswasta dan Prestasi bermasalah mendapatkan tingkat kelayakan kelas mikro.
- 5) jika jaminan SERTIFIKAT HM dengan pekerjaan PNS mendapatkan tingkat kelas kelayakan atas.
- 6) jika jaminan SERTIFIKAT HM dengan pekerjaan Swasta belum mendapatkan keputusan final, maka digunakan cabang ketiga yaitu cabang Prestasi dimana jaminan SERTIFIKAT HM dengan pekerjaan Swasta dan Prestasi baik mendapatkan tingkat kelayakan kelas atas dan sedang.

4.2. Perancangan ERD

Berikut ini merupakan bentuk Entity Relationship Diagram dari Sistem Informasi Klasifikasi Data Nasabah Menggunakan Metode Decision Tree dan Algoritma C4.5.

**Gambar 5. Rancangan ERD**

4.3. Desain Interface

Antar muka sistem informasi klasifikasi data nasabah menggunakan metode decision tree dan algoritma C4.5 berbasis web dapat dilihat dalam tampilan utama.

Logo

Home Menu Level User Logout

Pilihan Cabang

Prestasi
Jaminan
Pekerjaan

Tampilan Tabel Rekap kelayakan cabang

Gambar 6. Form Rekap Kelayakan

Logo

Home Menu Level User Logout

Update Data Kelayakan

No	Update Prestasi	Update Jaminan	Update Pekerjaan	Update Klas
1				
2				

Gambar 7. Form Update data Kelayakan

4.4. Hasil Implementasi Sistem

Atribut-atribut yang digunakan dalam penelitian ini adalah prestasi, jaminan dan pekerjaan, yang digunakan untuk menentukan pengklasifikasian nasabah kelas atas, kelas menengah dan kelas bawah :

LDSK
CPC Jawa Tengah

Home > Admin Web > Logout >

Tambah dan Update Data Sample Kelayakan

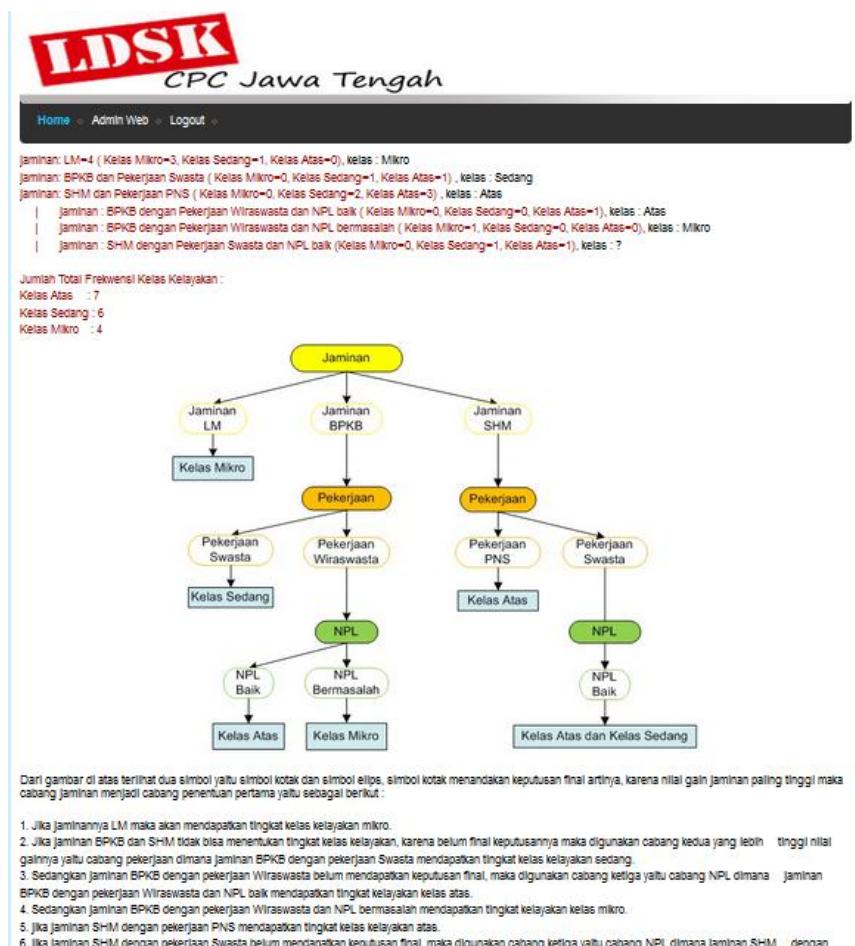
A Form isi Data Sample Kelayakan untuk ditampilkan pada menu Data Sample Kelayakan. Isi Data Sample Kelayakan dapat menggunakan tag html. Daftar di bawah form isian adalah daftar Data Sample Kelayakan yang telah tercatat dalam database. Masing-masing Data Sample Kelayakan dapat diedit ataupun dihapus.

Kode Nasabah :	00002
Nama Nasabah :	Soleman
Pilih NPL :	- Pilih NPL -
Pilihan Jaminan :	- Pilih Jaminan -
Pilihan Pekerjaan :	- Pilih Pekerjaan -
Pilihan Nominal Kelas :	- Pilih Kelas -
<input type="button" value="Input"/>	

Daftar Data Sample Nasabah

NO	NASABAH	NPL	JAMINAN	PEKERJAAN	KELAS	PILIHAN
1	Meidy Firman	Baik	SHM	PNS	Atas	edit hapus
2	Much. Rosul	Bermasalah	SHM	PNS	Atas	edit hapus
3	Arifin	Baik	BPKB	SWASTA	Sedang	edit hapus
4	Andi Ardiyanto	Baik	LM	SWASTA	Mikro	edit hapus
5	Anit Sofyan	Baik	BPKB	WIRASWASTA	Atas	edit hapus
6	Ad Herlambang	Bermasalah	BPKB	WIRASWASTA	Mikro	edit hapus
7	Bambang Ardiansyah	Baik	SHM	SWASTA	Sedang	edit hapus

Gambar 8. Tampilan Input Sampel Data Kelayakan



Gambar 9. Pohon Keputusan

4.5. Hasil Validasi Pakar Sistem

Tabel 5. Indikator Penilaian Pakar Sistem

No	Indikator	1	2	3	4
1	Kebenaran rancangan program			√	√
2	Program sudah menyediakan database untuk menampung data			√	√
3	Input data nasabah sesuai dengan rancangan			√	√
4	Program yang dibuat sesuai dengan kriteria yang digunakan			√	√
5	Proses kerja program sudah terlihat tahapan sistematisnya			√	√
6	Layout menu sudah sesuai dengan rancangan			√	√
7	Program yang dibuat terdapat menu perhitungan metode yang digunakan			√	√
8	Hasil sesuai dengan perhitungan			√	√
9	Sistem sudah memenuhi standart penggunaan Algoritma C4.5			√	√
10	Sistem klasifikasi nasabah decision tree sesuai dengan algoritma C4.5			√	√
	Jumlah Skor			9	28
	Jumlah Skor Total		37		

Keterangan :

1. tidak tepat / tidak menarik / tidak layak / tidak sesuai
2. kurang tepat / kurang menarik / kurang layak / kurang sesuai
3. tepat / menarik / layak / sesuai
4. sangat tepat / sangat menarik / sangat layak / sangat sesuai

Sehingga dapat dihitung nilai validasinya :

$$\mu = \frac{\sum x}{n} = \frac{37}{10} = 3,7$$

Keterangan :

- μ : nilai rata-rata
 $\sum x$: jumlah total skor validasi
 N : jumlah indikator penilaian

Berdasarkan perhitungan di atas, dapat disimpulkan bahwa hasil validasi dari pakar sistem menunjukkan nilai 3,7 berada diantara 3,26 – 4,00 yakni tergolong dalam kategori sangat valid.

5. KESIMPULAN

Setelah melakukan analisis, perancangan, implementasi beserta pengujian yang telah dilakukan, maka dapat disimpulkan bahwa aplikasi yang dibangun dapat membantu bagian Loan Doc and Safe Keeping menganalisa data nasabah untuk menentukan klasifikasi nasabah sehingga diharapkan bisa meminimalkan waktu seefisien mungkin dalam menentukan kelas nasabah. Jadi perancangan aplikasi data mining dan algoritma C4.5 ini telah sesuai dengan prosedur dan tujuan yang diharapkan. Hasil validasi dari pakar sistem menunjukkan nilai 3,7 berada diantara 3,26 – 4,00 yakni tergolong dalam kategori sangat valid.

DAFTAR PUSTAKA

- Kusrini**, 2009; " Perbandingan Metode Nearest Neighbor dan Algoritma C4.5 untuk Menganalisa Kemungkinan Pengunduran Diri Calon Mahasiswa di STIMIK Amikom Yogyakarta", Jurnal DASI Vol. 10 No.1 Maret 2009, ISSN: 1411-3201.
- Han, J & Kamber, M**, 2006; "Data Mining Concept and Techniques", San Francisco.
- Basuki, Achmad dan Syarif, Iwan**. 2003; "Modul Ajar Decision Tree", Surabaya : PENS-ITS.
- Sugiyono**. 2012; "Metode Penelitian Bisnis", Bandung : Alfabeta.

Borg, W.R. & Gall, M.D. Gall. (1989). *Educational Research: An Introduction, Fifth Edition*. New York: Longman.

Goller, 2000; "Automatic Document Classification: A Thorough Evaluation of Various Methods", Proceedings of International Symposium on information Theory and Its Application, pp. 145-162, USA.

Feldman, D.C. 2004; "Managing Individual Are Group. Behavioral in Organization", New York: McGraw Hill.