

Penerapan Metode C5.0 Untuk Pengelompokan Potensi Nasabah PT. Pegadaian Berdasarkan Pola Pembayaran Angsuran

Ika Berutu^{1*}

¹Fakultas Ilmu Komputer dan Teknologi Informasi, Program Studi Teknik Informatika, Universitas Budi Darma, Medan, Indonesia
Email: ^{1*}ikaaberutu24@gmail.com

Abstrak– Pengelompokan potensi nasabah sangat diperlukan untuk mengetahui nasabah yang potensial. Nasabah PT. Pegadaian sering terlambat melakukan pembayaran angsuran pinjaman dan sering lewat dari tanggal jatuh tempo yang telah ditentukan. Hal ini dapat menyebabkan kerugian pada PT. Pegadaian sendiri. Dalam penelitian ini akan menggunakan data mining dengan metode C5.0 menganalisa potensi nasabah berdasarkan pola pembayaran angsuran dengan cara pemilihan atribut yang akan diproses menggunakan information gain. Atribut dengan nilai information gain tertinggi akan dipilih sebagai parent bagi node selanjutnya. Data nasabah akan dianalisa untuk mendapatkan keputusan. Hal ini dapat dibuktikan bahwa perhitungan manual sama dengan perhitungan di aplikasi. Adapun hasil yang diperoleh merupakan pohon keputusan.

Kata Kunci: Nasabah, Data Mining dan Metode C5.0

Abstract–Classification of potential customers is necessary to identify potential customers. Customer of PT. Pegadaian is often late in making loan installment payments and is often past the predetermined due date. This can cause losses to PT. The pawn shop itself. In this study, data mining with the C5.0 method will be used to analyze potential customers based on installment payment patterns by selecting attributes that will be processed using information gain. The attribute with the highest information gain value will be chosen as the parent of the next node. Customer data will be analyzed to get a decision. It can be proven that manual calculations are the same as calculations in the application. The results obtained are a decision tree.

Keywords: Customer, Data Mining and C5.0 Method

1. PENDAHULUAN

Menurut Kamus Perbankan, Nasabah adalah orang atau badan yang mempunyai rekening simpanan atau pinjaman pada bank. Pada tahun 1998 melalui Undang-Undang Nomor 10 Tahun 1998 diintroduksilah rumusan masalah nasabah dalam pasal 1 angka 16, yaitu pihak yang menggunakan jasa bank. Potensi adalah sebuah kemampuan dasar yang dimiliki manusia yang sangat mungkin dikembangkan. Potensi nasabah berdasarkan pola pembayaran angsuran sangat diperlukan karena dengan itu kita dapat mengetahui potensial nasabah dalam melakukan pembayaran angsuran dengan waktu yang tepat.

Pegadaian adalah salah satu bentuk lembaga keuangan bukan bank di Indonesia yang mempunyai kegiatan membiayai kebutuhan masyarakat, baik itu bersifat produktif maupun konsumtif dengan menggunakan hukum gadai. Dalam Kitab Undang-Undang Hukum Perdata pasal 1150, gadai adalah suatu hak yang diperoleh pihak yang mempunyai piutang (pegadaian) atas suatu barang bergerak. Nasabah potensial adalah nasabah yang selalu tepat dalam melakukan pembayaran angsuran tanpa melewati jatuh tempo terakhir dan tidak merugikan perusahaan.

Data mining adalah proses yang mempekerjakan satu atau lebih teknik pembelajaran komputer (machine learning) untuk menganalisis dan mengekstraksi pengetahuan (knowledge) secara otomatis. Definisi lain diantaranya adalah pembelajaran berbasis induksi (induction-based learning) adalah proses pembentukan definisi-definisi konsep umum yang dilakukan dengan cara mengobservasi contoh-contoh spesifik dari konsep-konsep yang akan di pelajari. Knowledge Discovery in Databases (KDD) adalah penerapan metode saintifik pada data mining. Dalam konteks ini data mining merupakan satu langkah dari proses KDD.

Pada PT. Pegadaian nasabah banyak yang terlambat dalam melakukan pembayaran angsuran pinjaman. Hal tersebut dapat menimbulkan kerugian tersendiri bagi PT. Pegadaian karena tidak memaksimalkan kesempatan yang ada. Salah satu tantangan yang di hadapi PT. Pegadaian adalah keterlambatan nasabah melakukan pencicilan angsuran yang sudah jatuh tempo terakhir.

Oleh karena itu sangat diperlukan pengelompokan potensi nasabah berdasarkan berbagai kriteria seperti pola pembayaran angsuran, pekerjaan, status pernikahan, jumlah tanggungan, pendidikan dan penghasilan dan. Metode C5.0 merupakan salah satu metode dalam data mining yang termasuk dalam klasifikasi/pengelompokan berdasarkan kriteria yang berbasis numerik dengan output pohon keputusan.

Dalam penelitian ini penulis akan menggunakan metode C5.0 Berdasarkan jurnal dengan judul “Penerapan Data Mining dalam menganalisa pemberian pinjaman dengan menggunakan metode C5.0 disebutkan bahwa Untuk menganalisa kredit tersebut diperlukan suatu metode data mining dalam pemberian pinjaman salah satunya algoritma C5.0. Algoritma C5.0 digunakan untuk memberikan keputusan dalam pengajuan kredit. Decision tree merupakan salah satu metode klasifikasi yang menggunakan representasi struktur pohon (tree) dimana setiap node mempresentasikan atribut, cabangnya mempresentasikan nilai dari atribut dan daun mempresentasikan kelas. Node yang paling atas dari decision tree merupakan node klasifikasi yang paling populer digunakan.

2. METODOLOGI PENELITIAN

2.1 Data Mining

Data mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam database. *Data mining* adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terakut dari berbagai database besar [1].

2.2 Nasabah

Nasabah adalah orang atau badan hukum yang mempunyai rekening baik rekening simpanan atau pinjaman pada pihak bank, Sehingga nasabah merupakan orang yang biasa berhubungan dengan atau menjadi pelanggan bank [4]. Dengan kata lain nasabah adalah pihak atau orang yang menggunakan dan secara sengaja menjadi langganan bank yang di percayainya.

2.3 Metode C5.0

C5.0 menghasilkan tree dengan jumlah cabang per node bervariasi. C5.0 memerlukan variable kontinu sama dengan yang dilakukan oleh CART, tetapi untuk variable kategorika C5.0 memperlakukan nilai variable kategorikal sebagai splitter [6]. Strategi pengembangan decision tree dengan menggunakan metode C5.0 adalah sebagai berikut :

1. Pada tahap awal, tree digambarkan sebagai node tunggal yang mempresentasikan training set.
2. Jika sampel seluruhnya berisi kelas yang sama, maka node tersebut menjadi leaf dan diberi dengan kelas tersebut.
3. Jika tidak, algoritma dengan menggunakan ukuran berbasis entropy (information gain) akan memilih table predictor yang akan memisahkan record kedalam kelas-kelas individual. Variable tersebut menjadi variable tes atau keputusan pada node tersebut.
4. Cabang dikembangkan untuk tiap nilai yang diketahui dari variable tes, dan sampel di partisi berdasarkan cabang tersebut.
5. Algoritma menggunakan proses yang sama secara rekursif membentuk decision tree.
6. Partisi rekursif berakhir hanya ketika satu dari kondisi-kondisi berikut :
 - a. Seluruh record pada node tertentu memiliki kelas yang sama.
 - b. Tidak ada atribut yang tersisa pada record yang dapat dipartisi lebih lanjut. Dalam kasus ini suara mayoritas digunakan. Node tersebut menjadi leaf node dan dilabeli dengan kelas yang menjadi mayoritas dalam record yang ada.
 - c. Tidak ada record untuk cabang variable tes. Dalam kasus ini, leaf berbentuk dengan mayoritas kelas sebagai label record tersebut

Untuk menentukan akar dari pohon keputusan ditentukan oleh gain yang tertinggi, sebelum menemukan gain terdahulu menghitung entropy keseluruhan dan entropy dari setiap atribut seperti yang terlihat pada rumus dibawah [6]:

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

Setelah entropy diperoleh selanjutnya tentukan gain dari setiap atribut, dan mencari gain yang tertinggi untuk dijadikan akar dari cabang pohon keputusan seperti rumus dibawah ini :

$$\text{Gain}(S,A) = \text{Entropy}(S) - \sum_{i=1}^n |s_i| * \text{Entropy}(S_i) \dots\dots\dots (2)$$

3. HASIL DAN PEMBAHASAN

3.1 Analisa Masalah

Masalah yang akan di analisis adalah untuk pengelompokkan potensi nasabah berdasarkan pola pembayaran angsuran. Nasabah PT. Pegadaian sering terlambat dalam melakukan pembayaran angsuran. Hal ini dapat mengakibatkan kerugian tersendiri bagi PT. Pegadaian. Data selanjutnya akan dilakukan pra-proses untuk menghasilkan data kasus yang siap dibentuk untuk menjadi sebuah pohon keputusan.

Pada data mining digunakan metode C5.0. Metode C5.0 merupakan satu metode data mining yang khususnya diterapkan pada teknik decision tree. Dengan menggunakan klasifikasi metode C5.0 dapat membantu PT. Pegadaian upc kampung durian untuk pengelompokkan potensi nasabah. Dalam metode ini pemilihan atribut yang akan diproses menggunakan information gain. Dalam memilih atribut untuk pemecah obyek dalam beberapa kelas harus dipilih atribut yang menghasilkan information gain paling besar. Atribut dengan nilai information gain tertinggi akan dipilih sebagai parent bagi node selanjutnya.

Adapun analisa yang di lakukan di PT. Pegadaian upc kampung durian, dimana data-data nasabah dan data angsuran diambil pada tahun 2018. Dimana data-data nasabah dan angsuran yang digunakan untuk analisa ada 15 jenis data nasabah beserta atributnya.

3.1.1 Penerapan Metode C5.0

Metode C5.0 merupakan salah satu metode dalam data mining yang termasuk dalam klasifikasi/pengelompokan berdasarkan kriteria yang berbasis numerik dengan output pohon keputusan. Langkah-langkah metode decision tree C5.0 dalam melakukan proses data mining adalah sebagai berikut :

1. Mengelompokkan setiap variabel untuk mendapatkan atribut yang akan digunakan untuk menghitung entropy dari kasus agar mendapatkan akar dan cabang dari data angsuran nasabah. Dalam tahap ini juga tahap klasifikasi pada data mining dilakukan, dimana data tersebut dikelompokkan dan diklasifikasikan berdasarkan variabel-variabel yang telah dilakukan.

Table 1. Sampel Data yang akan digunakan

No	Nama Nasabah	Status Perkawinan	Pendidikan	Pekerjaan	Penghasilan	Jlh Tanggungan	Keterangan
1	Hotnida Esrawaty	Menikah	S1	Pegawai negeri	Sedang	Sedikit	Lancar
2	Helmina me pardede	Menikah	SMA	Wiraswasta	Sedang	Banyak	Lancar
3	Maruli Hutagulung	Menikah	SMA	Wiraswasta	Kecil	Sedikit	Lancar
4	Dewi Simanjuntak	Belum Menikah	S1	Pegawai negeri	Besar	Tidak Ada	Kurang Lancar
5	Dertina Simanullang	Belum Menikah	D3	Swasta	Besar	Tidak Ada	Lancar
6	Januar Yosua Purba	Menikah	SMA	Swasta	Sedang	Sedikit	Lancar
7	Anry Gumora Simamora	Menikah	S1	Pegawai negeri	Kecil	Banyak	Tidak Lancar
8	Hantus L Tobing	Menikah	SMA	Wiraswasta	Sedang	Sedikit	Tidak Lancar
9	Melindasari Napitupulu	Menikah	S1	Pegawai negeri	Kecil	Banyak	Tidak Lancar
10	Susanti	Belum Menikah	S1	Pegawai negeri	Besar	Tidak Ada	Lancar
11	Lia Novita Handayani	Menikah	D3	Wiraswasta	Kecil	Banyak	Kurang Lancar
12	Emil Murfi Siantury	Menikah	SMA	Wiraswasta	Besar	Sedikit	Kurang Lancar
13	Rostiawan Simorangkir	Menikah	SMA	Swasta	Kecil	Banyak	Tidak Lancar
14	Daud Hutabalian	Menikah	D3	Wiraswasta	Kecil	Banyak	Kurang Lancar
15	Hotnida Silalahi	Menikah	SMA	Wiraswasta	Sedang	Sedikit	Lancar

Dalam kasus yang tertera pada tabel 1. akan dibuat pohon keputusan untuk menentukan potensi nasabah berdasarkan pola pembayaran angsuran dengan melihat status perkawinan, pendidikan, pekerjaan, penghasilan dan jumlah tanggungan. Selanjutnya data tersebut akan diproses sesuai langkah-langkah membentuk pohon keputusan.

Untuk memilih atribut sebagai akar, didasarkan pada nilai gain tertinggi dari atribut- atribut yang ada. Untuk mendapatkan nilai gain, terlebih dahulu mencari nilai entropy. Dengan menggunakan dua persamaan berikut, seperti yang telah dijelaskan pada bab sebelumnya maka akan didapatkan nilai entropy dan gain yang digunakan sebagai akar dalam membuat pohon keputusan. Berikut ini merupakan rumus (1) untuk perhitungan nilai entropy, seperti pada persamaan satu berikut:

$$\text{Entropy (S)} = \sum_{i=0}^n - p_i * \log_2 p_i$$

Berikut ini merupakan perhitungan nilai entropy

1. Menghitung Entropy Total

Entropy Total adalah menghitung nilai Lancar, Kurang Lancar dan Tidak Lancar berdasarkan penilaian keseluruhan data berdasarkan tabel Data nasabah.

Total Nasabah : 15

Nilai atribut Lancar : 7

Nilai atribut Kurang Lancar : 4

Nilai atribut Tidak Lancar : 4

$$\text{Entropy (Total)}: \left(-\frac{7}{15} * \log_2 \left(\frac{7}{15} \right) \right) + \left(-\frac{4}{15} * \log_2 \left(\frac{4}{15} \right) \right) + \left(-\frac{4}{15} * \log_2 \left(\frac{4}{15} \right) \right)$$

Entropy (Total) : 1.530124

2. Menghitung Entropy Status perkawinan

- a. Atribut menikah

$$= \left(-\frac{5}{12} * \log_2 \left(\frac{5}{12} \right) \right) + \left(-\frac{2}{12} * \log_2 \left(\frac{2}{12} \right) \right) + \left(-\frac{4}{12} * \log_2 \left(\frac{4}{12} \right) \right)$$

$$= 1.485411$$

- b. Atribut belum menikah

$$= \left(-\frac{2}{3} * \log_2 \left(\frac{1}{3} \right) \right) + \left(-\frac{1}{3} * \log_2 \left(\frac{1}{3} \right) \right) + \left(-\frac{0}{3} * \log_2 \left(\frac{0}{3} \right) \right)$$

$$= 0.918294$$

3. Menghitung Entropy pendidikan

- a. Atribut S1

$$= \left(-\frac{2}{5} * \log_2 \left(\frac{2}{5} \right) \right) + \left(-\frac{1}{5} * \log_2 \left(\frac{1}{5} \right) \right) + \left(-\frac{2}{5} * \log_2 \left(\frac{2}{5} \right) \right)$$

$$= 1.521927$$

- b. Atribut SMA

$$= (-\frac{4}{7} * \log_2(\frac{4}{7})) + (-\frac{1}{7} * \log_2(\frac{1}{7})) + (-\frac{2}{7} * \log_2(\frac{2}{7}))$$

$$= 1.378781$$
- c. Atribut D3

$$= (-\frac{1}{3} * \log_2(\frac{1}{3})) + (-\frac{2}{3} * \log_2(\frac{2}{3})) + (-\frac{0}{3} * \log_2(\frac{0}{3}))$$

$$= 0.918294$$
4. Menghitung Entropy pekerjaan
- a. Atribut Pegawai Negeri

$$= (-\frac{2}{5} * \log_2(\frac{2}{5})) + (-\frac{1}{5} * \log_2(\frac{1}{5})) + (-\frac{2}{5} * \log_2(\frac{2}{5}))$$

$$= 1.521927$$
- b. Atribut Wiraswasta

$$= (-\frac{3}{7} * \log_2(\frac{3}{7})) + (-\frac{3}{7} * \log_2(\frac{3}{7})) + (-\frac{1}{7} * \log_2(\frac{1}{7}))$$

$$= 1.448814$$
- c. Atribut Swasta

$$= (-\frac{2}{3} * \log_2(\frac{2}{3})) + (-\frac{0}{3} * \log_2(\frac{0}{3})) + (-\frac{1}{3} * \log_2(\frac{1}{3}))$$

$$= 0.918294$$
5. Menghitung Entropy penghasilan
- a. Atribut Sedang

$$= (-\frac{4}{5} * \log_2(\frac{4}{5})) + (-\frac{0}{5} * \log_2(\frac{0}{5})) + (-\frac{1}{5} * \log_2(\frac{1}{5}))$$

$$= 0.721927$$
- b. Atribut Kecil

$$= (-\frac{1}{6} * \log_2(\frac{1}{6})) + (-\frac{2}{6} * \log_2(\frac{2}{6})) + (-\frac{3}{6} * \log_2(\frac{3}{6}))$$

$$= 1.459147$$
- c. Atribut Besar

$$= (-\frac{2}{4} * \log_2(\frac{2}{4})) + (-\frac{2}{4} * \log_2(\frac{2}{4})) + (-\frac{0}{4} * \log_2(\frac{0}{4}))$$

$$= 1$$
6. Menghitung Entropy jumlah tanggungan
- a. Atribut Sedikit

$$= (-\frac{4}{6} * \log_2(\frac{4}{6})) + (-\frac{1}{6} * \log_2(\frac{1}{6})) + (-\frac{1}{6} * \log_2(\frac{1}{6}))$$

$$= 1.251628$$
- b. Atribut Banyak

$$= (-\frac{1}{6} * \log_2(\frac{1}{6})) + (-\frac{2}{6} * \log_2(\frac{2}{6})) + (-\frac{3}{6} * \log_2(\frac{3}{6}))$$

$$= 1.459147$$
- c. Atribut Tidak ada

$$= (-\frac{2}{3} * \log_2(\frac{2}{3})) + (-\frac{1}{3} * \log_2(\frac{1}{3})) + (-\frac{0}{3} * \log_2(\frac{0}{3}))$$

$$= 0.918294$$

Sementara itu, untuk mencari nilai gain dapat dilihat pada persamaan berikut:

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

Berikut ini merupakan perhitungan nilai gain :

- Gain status perkawinan
 Berikut adalah nilai gain dari setiap kriteria:

$$= 1.530124 - ((\frac{12}{15}) * -1.554691) + ((\frac{3}{15}) * 0.918306)$$

$$= 1.530137 - 1.371986$$

$$= 0.158138$$
- Gain Pendidikan
 Berikut adalah nilai gain dari setiap kriteria:

$$= 1.530124 - ((\frac{5}{15}) * -1.52192) + ((\frac{7}{15}) * 1.378741) + ((\frac{3}{15}) * 0.918289)$$

$$= 1.530124 - 1.334398$$

$$= 0.195726$$
- Gain Pekerjaan
 Berikut adalah nilai gain dari setiap kriteria:

$$= 1.530124 - ((\frac{5}{15}) * 1.52192) + ((\frac{7}{15}) * 1.448812) + ((\frac{3}{15}) * 0.918289)$$

$$= 1.530124 - 1.36708$$

$$= 0.163044$$

4. Gain Penghasilan

Berikut adalah nilai gain dari setiap kriteria:

$$= 1.530124 - ((\frac{5}{15}) * 0.72192) + ((\frac{6}{15}) * 1.459123) + ((\frac{4}{15}) * 1)$$

$$= 1.530124 - 1.090966$$

$$= 0.439158$$

5. Gain Jumlah Tanggungan

Berikut adalah nilai gain dari setiap kriteria:

$$= 1.530124 - ((\frac{6}{15}) * -0.820784) + ((\frac{6}{15}) * 1.459123) + ((\frac{3}{15}) * 0.918289)$$

$$= 1.530124 - 1.267967$$

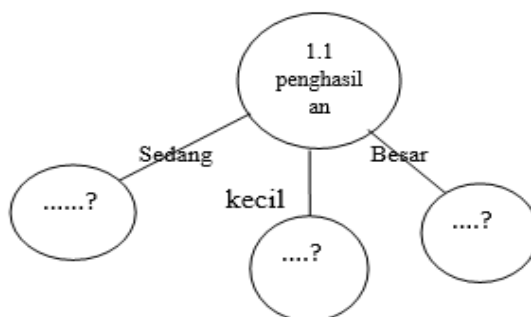
$$= 0.262157$$

Setelah di dapatkan nilai entropy dan gain dari sampel data yang dimiliki, berikut ini adalah rekapitulasi perhitungan nilai entropy dan gainnya

Tabel 2. Rekapitulasi Hasil

Node		Keterangan	Jlh Kasus (s)	Lancar	Kurang Lancar	Tidak Lancar	Entropy	Gain
1	TOTAL		15	7	4	4	1.530.124	
	Status perkawinan							0.158138
		Menikah	12	5	3	4	1.485.411	
		Belum Menikah	3	3	1	0	0.918294	
	Pendidikan							
		S1	5	2	1	2	1.521.927	0.195726
		SMA	7	4	1	2	1.378.781	
		D3	3	1	2	0	0.918294	
	Pekerjaan							0.163044
		Pegawai negeri	5	2	1	2	1.521.927	
		Wiraswasta	7	3	3	1	1.448.814	
		Swasta	3	2	0	1	0.918294	
	Penghasilan							0.439158
		Sedang	5	4	0	1	0.721927	
		Kecil	6	1	2	3	1.459.147	
		Besar	4	2	2	0	1	
	jlh tanggungan							0.262157
		Sedikit	6	4	1	1	1.251.628	
		Banyak	6	1	2	3	1.459.147	
		Tidak ada	3	2	1	0	0.918294	

Dari hasil tabel 2. dapat diketahui bahwa atribut dengan nilai gain tertinggi adalah penghasilan yaitu sebesar 0.439158. Dengan demikian penghasilan dapat menjadi node akar. Ada tiga nilai atribut dari penghasilan yaitu sedang, kecil, besar. Dari ketiga nilai atribut tersebut, nilai ketiga atribut tersebut belum mendapatkan keputusan sehingga perlu dilakukan perhitungan. Setelah dilakukan perhitungan maka terbentuklah pohon keputusan sementara seperti pada gambar 1.



Gambar 1. Node

Pada node 1. atribut dengan nilai gain tertinggi adalah penghasilan sehingga penghasilan menjadi node akar. Ada tiga nilai atribut penghasilan yaitu sedang, kecil dan besar belum ada hasil keputusan, sehingga masih perlu dilakukan perhitungan. Pohon keputusan di atas belum terlihat keputusan yang dominan dari setiap keterangan yang di pilih. Maka kita harus mencari kembali nilai entropy dan gain dari setiap atribut (kriteria) penghasilan =kecil

1. Nilai entropy

Berikut ini adalah tabel penyelesaiannya.

Tabel 3. Sampel Data Yang Di Uji Ulang (Kriteria Penghasilan)

Kriteria	Atribut	jumlah	Lancar	Kurang lancar	Tidak Lancar
		Kasus			
Penghasilan					
	Kecil	6	1	2	3

Setelah itu kita hitung nilai entropy dari atribut Penghasilan = Kecil yang memiliki jumlah kasus “6”, karena atribut kecil memiliki jumlah kasus paling tinggi seperti terlihat pada tabel di bawah ini.

Tabel 4. Atribut Penghasilan

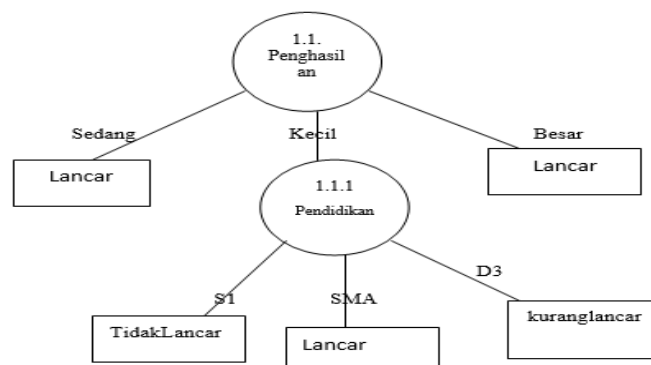
No	Nama Nasabah	Status Perkawinan	Pendidikan	Pekerjaan	Penghasilan	Jlh Tanggungan	Keterangan
3	Maruli Hutagulung	Menikah	SMA	Wiraswasta	Kecil	Sedikit	Lancar
7	Anry Gumora Simamora	Menikah	S1	Pegawai negeri	Kecil	Banyak	Tidak Lancar
9	Melindasari Napitupulu	Menikah	S1	Pegawai negeri	Kecil	Banyak	Tidak Lancar
11	Lia Novita Handayani	Menikah	D3	Wiraswasta	Kecil	Banyak	Kurang Lancar
13	Rostiawan Simorangkir	Menikah	SMA	Swasta	Kecil	Banyak	Tidak Lancar
14	Daud Hutabalian	Menikah	D3	Wiraswasta	Kecil	Banyak	Kurang Lancar

Langkah selanjutnya menghitung nilainya, berikut ini adalah rekapitulasi nilai entropy dan gainnya. Tabel ini di buat berdasarkan atribut yang akan di hitung yaitu atribut kecil.

Tabel 5. Hasil rekapitulasi atribut penghasilan

Node	Keterangan	Jlh Kasus	Lancar	Kurang Lancar	Tidak Lancar	Entropy	Gain
		(s)					
Penghasilan							
	Kecil	6	1	2	3	1.459.145	
status perkawinan							0
	Menikah	6	1	2	3	1.459.145	
Pendidikan							1.125.812
	S1	2	0	0	2	0	
	SMA	2	1	0	1	0	
	D3	2	0	2	0	0	
Pekerjaan							0
	Pegawai negeri	2	0	2	0	0	
	Wiraswasta	2	0	2	0	0	
	Swasta	1	0	0	1	0	
jlh tanggungan							0
	Sedikit	1	1	0	0	0	
	Banyak	5	0	2	3	0.97095	

Berdasarkan tabel di atas terlihat bahwasanya atribut = Pendidikan memiliki nilai gain tertinggi, maka untuk root selanjutnya pada pohon keputusannya dapat terlihat pada gambar pohon (tree) berikut ini:



Gambar 2. Node Keputusan

Dengan memperhatikan pohon keputusan pada gambar 2., diketahui bahwa semua kasus sudah masuk dalam kelas. Dengan demikian, pohon keputusan pada gambar 2. merupakan keputusan terakhir yang terbentuk. Maka basis pengetahuan atau rule yang terbentuk yaitu:

1. Jika penghasilannya = sedang, maka hasil=lancar
2. Jika penghasilannya = kecil, pendidikan = S1 maka hasil = tidak lancar
3. Jika penghasilannya = besar, maka hasil = lancar
4. Jika penghasilannya = kecil, pendidikan = SMA maka hasil = lancar
5. Jika penghasilannya = kecil, pendidikan = D3 maka hasil = kurang lancar

Jika penghasilannya sedang maka hasilnya lancar. sedangkan jika penghasilannya kecil dan pendidikannya S1 maka hasilnya tidak lancar. Kemudian jika penghasilannya besar maka hasilnya lancar. Jika penghasilannya kecil, pendidikannya SMA maka hasilnya lancar sedangkan jika penghasilannya kecil, pendidikannya D3 maka hasilnya kurang lancar.

3.2. Pengujian

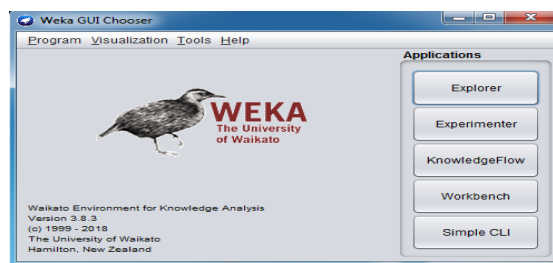
Pengujian ini dilakukan terhadap sampel contoh kasus yang akan diuji. Pengambilan contoh uji ini diambil dari beberapa kriteria dari sampel data yang digunakan. Pengujian ini dilakukan untuk menentukan apakah sampel tersebut mempunyai hasil keputusan yang sama.

Dalam pengujian data mining untuk Pengelompokan potensi nasabah berdasarkan pola pembayaran angsuran ini membutuhkan perangkat keras (hardware) dan perangkat lunak (software). Adapun hardware dan software yang akan dibutuhkan adalah sebagai berikut:

1. Perangkat Keras Minimum
 - a. Processor Intel Core I3
 - b. Memori 2 GB
 - c. Hardisk 500 GB
 - d. Monitor 12 Inchi
 - e. Keyboard
2. Perangkat Lunak Minimum
 - a. Sistem Operasi Windows 7 Ultimate 64 bit
 - b. Weka 3.8
 - c. Java
 - d. Microsoft Office Excel 2007
 - e. e. Notepad ++

Weka merupakan aplikasi yang dibuat dari bahasa pemrograman java yang dapat digunakan untuk membantu pekerjaan data mining (penggalian data). Weka berisi beragam jenis algoritma yang dapat digunakan untuk memproses dataset secara langsung atau bisa juga dipanggil melalui kode bahasa java.

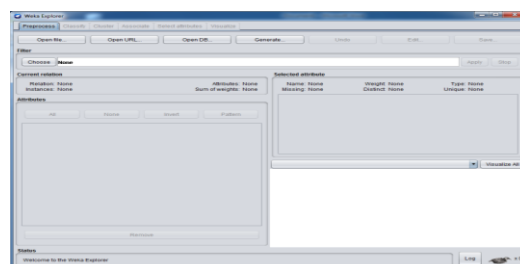
Weka berisi peralatan seperti pre-processing, classification, regression, clustering, association rules dan visualization. Weka dapat juga digunakan untuk memproses big data dan dikembangkan guna memenuhi skema machine learning (ML). Weka bersifat open source dibawah lisensi GNU (General Public License).



Gambar 3. Tampilan utama aplikasi weka

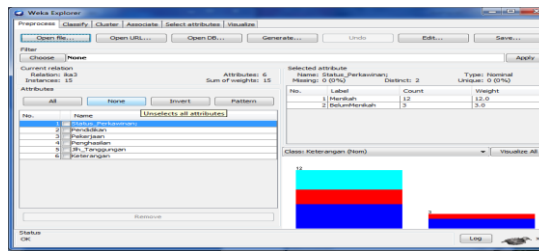
Setelah masuk ke halaman utama aplikasi Weka, pertama kali untuk mengubah database data nasabah (format csv) ke bentuk arf dengan cara :

1. Pilih menu explorer pada aplikasi tersebut. Lalu akan muncul tampilan pada gambar 4.



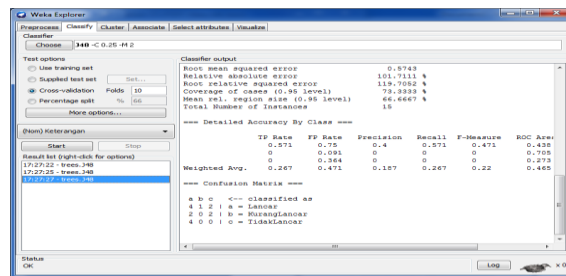
Gambar 4. Weka Explorer

- Kemudian buka database data nasabah (format csv) dengan meng-klik open file. Setelah database dibuka maka akan muncul tampilan sebagai berikut.



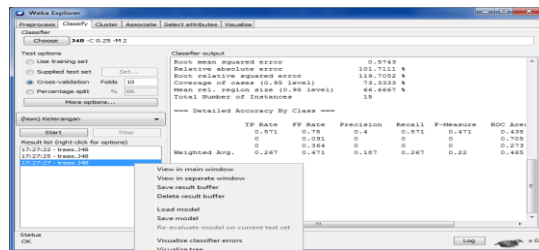
Gambar 5. Tampilan Database di aplikasi weka

Kemudian pilih classify pada menu bar untuk memilih algoritma yang akan digunakan seperti pada gambar dibawah ini.



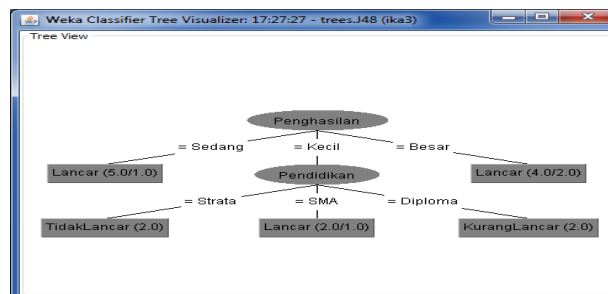
Gambar 6. Tampilan Classify Output

- Pada result list, klik kanan algoritma untuk menampilkan classify kemudian double click pada visualize tree.



Gambar 7. Tampilan Classify Output

Kemudian klik start, dan pilih visualize tree maka akan muncul hasil seperti gambar di bawah ini.



Gambar 8. Hasil keputusan

4. KESIMPULAN

Dalam penelitian skripsi ini dalam penerapan metode C5.0 untuk pengelompokan potensi nasabah PT. Pegadaian berdasarkan pola pembayaran angsuran, maka didapatkan kesimpulan mengelompokkan potensi nasabah berdasarkan pola pembayaran angsuran berdasarkan berbagai kriteria seperti pola pembayaran angsuran, pekerjaan, status perkawinan, jumlah tanggungan, pendidikan, dan penghasilan. Cara menerapkan metode C5.0 untuk pengelompokan potensi nasabah berdasarkan pola pembayaran agsuran dengan cara pemilihan atribut yang akan diproses menggunakan information gain. Dalam memilih atribut untuk pemecah obyek dalam beberapa kelas harus dipilih atribut yang menghasilkan information gain paling besar. Atribut dengan nilai information gain tertinggi akan dipilih sebagai parent bagi node selanjutnya

REFERENCES

- [1] Kusriani and Emha Taufiq Luthfi, *Algoritma Data Mining*, 1 ed., Theresia Ari Prabawanti, Ed. Yogyakarta: ANDI, 2009.
- [2] Finn Lee S and Juan Santana, *Data Mining: Meramalkan Bisnis Perusahaan*, Feri Sulianta and Dominikus Juju, Eds. Jakarta: PT. Elex Media Komputindo, 2010.
- [3] M.Pd.I, Arinda Firdianti, *Implementasi Manajemen Berbasis Sekolah Dalam Meningkatkan Prestasi Belajar Siswa*, Elin Wiji Astuti, Ed. Yogyakarta: CV. GRE PUBLISHING, 2018.
- [4] Al Arief M.Nur Rianto, *Dasar-Dasar Pemasaran Bank Syariah*. Bandung: Alfabeta, 2010.
- [5] Bill Stinnet, *Think Like Your Customer*. Jakarta: Alex Media Komputindo, 2008.
- [6] M.Kom, Febri Hadi. S.kom, "PENERAPAN DATA MINING DALAM MENGANALISA PEMBERIAN PINJAMANA DENGAN MENGGUNAKAN METODE ALGORITMA C5.0," UPI YPTK Jurnal KomTekInfo, vol. 4, no. 2356-0010, p. 216, Desember 2017.
- [7] Edhy Sutanta, *Pengantar Teknologi Informasi*. Yogyakarta: GRAHA ILMU, 2005.
- [8] Erdi Susanto, *Data Mining Menggunakan Weka*. Jakarta, 2012.
- [9] Adi Kusrianto, *Memfaatkan Formula dan Fungsi Microsoft Office Excel 2007*. Jakarta: PT.Elex Media Komputindo, 2007.
- [10] Jarot S, Shenia A, and Sudarman S, *Cara Mudah Menguasai Ms. Excel 2007 dalam seminggu*. Jakarta : media kita, 2010.
- [11] D. P. Utomo and Mesran, "Analisis Komparasi Metode Klasifikasi Data Mining dan Reduksi Atribut Pada Data Set Penyakit Jantung," J. MEDIA Inform. BUDIDARMA, vol. 4, no. 2, pp. 437–444, 2020.
- [12] D. P. Utomo, P. Sirait, and R. Yunis, "Reduksi Atribut Pada Dataset Penyakit Jantung dan Klasifikasi Menggunakan Algoritma C5. 0," J. MEDIA Inform. BUDIDARMA, vol. 4, no. 4, pp. 994–1006, 2020.
- [13] B. S. Pranata and D. P. Utomo, "Penerapan Data Mining Algoritma FP-Growth Untuk Persediaan Sparepart Pada Bengkel Motor (Study Kasus Bengkel Sinar Service)," Bull. Inf. Technol., vol. 1, no. 2, pp. 83–91, 2020.
- [14] D. P. Utomo and B. Purba, "Penerapan Datamining pada Data Gempa Bumi Terhadap Potensi Tsunami di Indonesia," Pros. Semin. Nas. Ris. Inf. Sci., vol. 1, no. 1, pp. 846–853, 2019.
- [15] R. Amelia and D. P. Utomo, "ANALISA POLA PEMESANAN PRODUK MODERN TRADE INDEPENDENT DENGAN MENEREPAKAN ALGORITMA FP. GROWTH (STUDI KASUS: PT. ADAM DANI LESTARI)," KOMIK (Konferensi Nas. Teknol. Inf. dan Komputer), vol. 3, no. 1, pp. 416–423, 2019.