

INFORMATION GAIN UNTUK MENGETAHUI PENGARUH ATRIBUT TERHADAP KLASIFIKASI PERSETUJUAN KREDIT

Much. Rifqi Maulana, M. Adib Al Karomi
Program Studi Sistem Informasi
STMIK Widya Pratama
Jl. Patriot 25 Pekalongan
Telp (0285) 427816
Email: rifq_13@yahoo.co.id , adib.comp@gmail.com

ABSTRAK

KREDIT MERUPAKAN SEBUAH PERILAKU EKONOMI YANG BANYAK DILAKUKAN OLEH MASYARAKAT MODERN. PERSETUJUAN KREDIT DAPAT MEMBANTU MASYARAKAT UNTUK MENDAPATKAN PINJAMAN SEJUMLAH UANG GUNA MEMBELI SUATU BARANG YANG UMUMNYA TIDAK DAPAT TERJANGKAU DENGAN PEMBAYARAN KONTAN. BANYAKNYA KRITERIA YANG DIAJUKAN OLEH PIHAK BANK ATAU PIHAK PENYEDIA DANA PADA UMUMNYA BERFUNGSI UNTUK MENYARING DATA NASABAH YANG NANTINYA DINYATAKAN LAYAK DIBERIKAN KREDIT ATAUPUN TIDAK. SEMAKIN BANYAK KRITERIA ATAU ATRIBUT YANG DIGUNAKAN TIDAK AKAN MENJAMIN KEAKURATAN KLASIFIKASI PERSETUJUAN KREDIT. BEBERAPA ATRIBUT YANG TIDAK BERPENGARUH JUSTRU AKAN MEMBUAT HASIL KLASIFIKASI MENJADI KURANG AKURAT. USAHA UNTUK MENGETAHUI BERAPA BESAR PENGARUH SEBUAH ATRIBUT DALAM KLASIFIKASI BANYAK DILAKUKAN OLEH PENELITI. SALAH SATUNYA MENGGUNAKAN TEKNIK DATA MINING. DALAM TAHAP *PRE PROCESSING* DATA MINING SALAH SATU YANG BANYAK DIGUNAKAN ADALAH SELEKSI FITUR. BEBERPA TEKNIK SELEKSI FITUR BERHASIL UNTUK MENINGKATKAN AKURASI KLASIFIKASI DATA MINING. DALAM PENELITIAN INI AKAN DILAKUKAN SELEKSI FITUR MENGGUNAKAN ALGORITMA *INFORMATION GAIN* UNTUK MENGETAHUI PENGARUH ATRIBUT TERHADAP KLASIFIKASI PERSETUJUAN KREDIT.

Kata kunci: Atribut data, *Information gain*, Persetujuan kredit



1. Pendahuluan

1.1 Latar Belakang

Dalam dunia perbankan istilah kredit sangat tidak asing terdengar. Kredit itu sendiri merupakan proses peminjaman sejumlah uang atau barang dengan nilai tertentu dan dibayar atau dilunasi dengan pembayaran secara berkala sampai dengan periode dan jumlah tertentu sesuai perjanjian. Dalam kredit ini pula dikenakan biaya bunga atau penambahan jumlah pembayaran sesuai dengan ketentuan pihak pemberi pinjaman. Sebagian nasabah melakukan pembayaran sesuai dengan ketentuan dan sebagian nasabah yang lain tidak melakukan pembayaran sesuai dengan ketentuan bahkan sebagian yang lain tidak melakukan pembayaran atas pinjaman yang telah diberikan. Dewasa ini banyak bermunculan aplikasi persetujuan kredit untuk menentukan kelayakan nasabah mendapatkan kredit dengan menggunakan berbagai metode komputasi.

Beberapa aplikasi tersebut antara lain menjadi sebuah sistem pendukung keputusan yang digunakan oleh pihak pemberi pinjaman guna memilih nasabah yang layak diberikan pinjaman. Salah satunya adalah SPK pada PT.BPR Artamanunggal Abadi Mranggen (Suhari, Sukur, and Eniyati 2009). SPK sejenis juga pernah dibuat dengan studi kasus Bank Muamalat cabang Yogyakarta (Dwi Cahyani 2013). Metode data mining juga dapat digunakan dalam klasifikasi kelayakan pemberian kredit (Ginjar Mabrur and Lubis 2012). Dalam prakteknya analisis data pembayaran kredit nasabah bank dengan memanfaatkan teknik data mining juga banyak dilakukan (Melissa and Oetama 2013). Selain dari itu banyak juga peneliti yang menggunakan metode seperti halnya *Analytical Hierarchy Process (AHP)* untuk SPK pemberian kredit pada nasabah bank (Azwany 2010). Tidak hanya itu saja, metode Bayes yang banyak digunakan untuk teknik data mining juga pernah digunakan untuk SPK sejenis (Zahid 2013). Model data mining juga dapat diimplementasikan untuk pembiayaan mobil (Wachid K). Dengan menggunakan metode *Decision Tree* (W, Pratikto, and Vivianne 2009) *Simple Additive Weighting* (Wahyu Oktaputra and Noersasongko 2014) serta metode *Scoring Sistem* (Wasana 2010).

Atribut dalam sebuah dataset dapat mempengaruhi hasil dari klasifikasi sebuah metode menggunakan teknik data mining (Han and Kamber 2006). Semakin banyak atribut yang relevan dapat meningkatkan akurasi algoritma data mining (Maimoon and Rokach 2010). Sebaliknya jika dalam dataset atribut yang digunakan banyak yang tidak relevan maka hasil klasifikasi juga akan semakin diragukan (Alpaydin 2010). Model pemilihan atribut dalam data mining sering juga disebut dengan *feature selection*. Dalam perkembangannya banyak muncul metode *feature selection* yang digunakan untuk *pre processing*. Salah satu metode *feature selection* yang banyak digunakan adalah *information gain* (Azhagusundari and Thanamani 2013). Metode *information gain* ini baik digunakan khususnya untuk dataset berdimensi tinggi (Koprinska 2010). Dalam penelitian ini akan digunakan algoritma *feature selection* untuk mengetahui pengaruh atribut dalam klasifikasi persetujuan kredit.

1.2 Landasan Teori

1.2.1 Data Mining

Data mining merupakan sebuah proses ekstraksi untuk mendapatkan suatu informasi yang sebelumnya tidak diketahui dari sebuah data (Witten, Frank, and Hall 2011). Data mining dapat menganalisa kasus lama untuk menemukan pola dari data dengan menggunakan teknik pengenalan pola seperti statistik dan matematika (Larose 2005). Data Mining atau sering juga disebut *Knowledge Discovery in Database* (KDD) adalah sebuah bidang ilmu yang banyak membahas tentang pola sebuah data. Serangkaian proses guna mendapatkan pengetahuan atau pola dari kumpulan data disebut dengan data mining (Ian H Witten. Eibe Frank. Mark A Hall 2011). Sebuah data yang besar bisa saja tidak berguna dan hanya akan menjadi sampah bila kita tidak dapat memanfaatkannya. Data mining menjawab masalah ini dengan menganalisa data yang besar tersebut kemudian membuat sebuah aturan, pola, ataupun model tertentu untuk mengenali data baru yang tidak berada dalam baris data yang tersimpan (Prasetyo 2012). Data mining merupakan kegiatan yang meliputi pengumpulan dan pemakaian data historis untuk menemukan keteraturan, pola atau hubungan dalam set data (Santosa 2007).

Dari beberapa pendapat pakar diatas, dapat disimpulkan bahwa data mining merupakan sebuah disiplin ilmu yang menggunakan data lampau sebagai patokan untuk menemukan pengetahuan baru.

1.2.2 Seleksi Fitur

Seleksi fitur merupakan suatu proses pemilihan atribut yang dianggap relevan dalam proses data mining. Banyaknya atribut akan mempengaruhi proses komputasi dan bahkan jika banyak atribut yang tidak relevan digunakan dalam sebuah proses klasifikasi maka dapat juga mempengaruhi hasil akurasi (Witten, Frank, and Hall 2011). Proses seleksi fitur sebenarnya adalah membuang atribut yang tidak relevan dan berlebihan. Proses pengurangan dimensi data ini juga memungkinkan algoritma klasifikasi untuk bekerja lebih cepat dan efektif serta memungkinkan peningkatan akurasi suatu algoritma.

1.2.3 Information Gain

Information gain merupakan salah satu metode seleksi fitur yang banyak dipakai oleh peneliti untuk menentukan batas dari kepentingan sebuah atribut (Deng and Runger 2012) (Azhagusundari and Thanamani 2013) (Novakovic 2010). Nilai *information gain* diperoleh dari nilai *entropy* sebelum pemisahan dikurangi dengan nilai *entropy* setelah pemisahan. Pengukuran nilai ini hanya digunakan sebagai tahap awal untuk penentuan atribut yang nantinya akan digunakan atau dibuang. Atribut yang memenuhi kriteria pembobotan yang nantinya akan digunakan dalam proses klasifikasi sebuah algoritma. Pemilihan fitur dengan *information gain* dilakukan dalam 3 tahapan, yaitu:

1. Menghitung nilai *information gain* untuk setiap atribut dalam dataset original.
2. Tentukan batas (*threshold*) yang diinginkan. Hal ini akan memungkinkan atribut yang berbobot sama dengan batas atau lebih besar akan dipertahankan serta membuang atribut yang berada dibawah batas.
3. Dataset diperbaiki dengan pengurangan atribut.

Pengukuran atribut ini pertama kali dipelopori oleh Claude Shannon didalam teori informasi (Gallager and Fellow 2001) serta dituliskan sebagai:

$$info(D) = -\sum_{i=1}^m p_i \log_2(p_i)$$

Keterangan:

- D : Himpunan kasus
- M : Jumlah partisi D
- p_i : Proporsi dari D_i terhadap D

Sedangkan p_i merupakan probabilitas sebuah *tuple* pada D yang masuk kedalam kelas C_i dan diestimasi dengan $|C_i, D| / |D|$. Fungsi log dalam hal ini digunakan log berbasis 2 karena informasi dikodekan berbasis bit. Perhitungan nilai *entropy* setelah pemisahan dapat dilakukan dengan menggunakan rumus berikut:

$$info_A(D) = -\sum_{j=1}^v \frac{|D_j|}{|D|} \times I(D_j)$$

Keterangan:

- D : Himpunan kasus
- A : Atribut
- v : Jumlah partisi atribut A
- $|D_j|$: Jumlah kasus pada partisi ke j
- $|D|$: Jumlah kasus dalam D
- $I(D_j)$: Total *entropy* dalam partisi

Sedangkan untuk mencari *information gain* atribut A dapat digunakan rumus berikut:

$$Gain(A) = I(D) - I(A)$$

Keterangan:

- Gain(A) : *Information* atribut A
- $I(D)$: Total *entropy*
- $I(A)$: *entropy* A

Gain (A) merupakan reduksi yang diharapkan di dalam *entropy* yang disebabkan oleh pengenalan nilai atribut dari A. Atribut yang memiliki nilai *information gain* terbesar selanjutnya dipilih sebagai uji atribut himpunan S. Lalu suatu simpul dibuat dan diberi label sesuai dengan atribut tersebut, kemudian cabang-cabang dibuat untuk masing-masing nilai atribut yang lain.

2. METODE PENELITIAN

Metode penelitian pada penelitian ini adalah penelitian eksperimen. Data nasabah kredit diambil dari sebuah perusahaan pembiayaan. Dalam dataset sebenarnya

terdapat 16 atribut yang memiliki berbagai varian nilai. Kemudian tahap berikutnya keseluruhan atribut akan dihitung dengan menggunakan *information gain* untuk mengetahui kepentingan dari atribut terhadap klasifikasi. Dan pada akhirnya setelah diketahui kepentingan dari semua atribut terhadap hasil klasifikasi maka atribut yang kiranya tidak perlu dapat dihilangkan dari dataset dan dataset diperbaharui. Metode pengumpulan data sampai dengan perhitungan secara lebih terperinci akan disampaikan dalam sub bab berikut:

2.1 Metode Pengumpulan Data

Tahapan pertama yang dilakukan dalam penelitian ini adalah tahap pengumpulan data. Seperti yang telah dijelaskan sebelumnya bahwa data yang akan digunakan adalah data nasabah kredit dalam sebuah perusahaan pembiayaan. Dalam dataset yang ada terdapat keseluruhan 766 *record* pelanggan yang pernah menerima pinjaman kredit dengan 210 dari pelanggan tersebut tergolong dalam nasabah dengan kredit lancar serta sisanya 556 tergolong dalam kredit macet.

Dalam dataset tersebut terdapat 16 atribut data dengan satu diantaranya menjadi atribut id yaitu nama nasabah serta satu menjadi atribut label atau sering juga disebut dengan atribut tujuan yaitu status kredit. Artinya dalam dataset terdapat 14 atribut reguler yang nantinya akan digunakan untuk klasifikasi dengan menggunakan algoritma C4.5. atribut tersebut antara lain: jenis kelamin, umur, jumlah pinjaman, jangka waktu pembayaran, angsuran tiap bulan, tipe pinjaman, jenis pinjaman, bi sektor ekonomi, col, bi golongan debitur, bi golongan penjamin, saldo nominatif, tunggakan pokok serta tunggakan bunga. Secara lebih terperinci meta data dari dataset persetujuan kredit dapat dilihat pada tabel 1 berikut:

Tabel 1 Meta Data Data Persetujuan Kredit

	Atribut	Type	Range
1	Nama	Polynomial (id)	~
2	Jenis Kelamin	Binominal	L, P
3	Umur	numerik	~
4	Jumlah Pinjaman	numerik	~
5	Jangka Waktu	Polynomial	~
6	Angsuran Perbulan	numerik	~
7	Type Pinjaman	nominal	100
8	Jenis Pinjaman	Polynomial	301, 302, 303, 304, 305
9	BI Sektor Ekonomi	Polynomial	6000, 8000, 9990
10	col	Binominal	1, 2
11	BI Golongan Debitur	Polynomial	834, 874, 876
12	BI Golongan Penjamin	Polynomial	000, 800, 835, 874, 875
13	Saldo Nominatif	numerik	~

14	Tunggakan Pokok	numerik	~
15	Tunggakan Bunga	numerik	~
16	Status Kredit	Binominal (label)	Lancar, Macet

2.2 Tahap Eksperimen

Tahap eksperimen dilakukan dengan menggunakan *tools* Rapid miner. Proses seleksi fitur dilakukan dengan hanya menggunakan fasilitas yang sudah ada dalam *tools* tersebut. Dalam tahap eksperimen ini tidak dilakukan koding ataupun membuat sintak tertentu untuk menghitung algoritma. Dikarenakan dalam *tools* Rapid Miner sudah tersedia banyak algoritma data mining yang dalam penggunaannya hanya tinggal *drag and drop*.

Dalam tahapan ini akan dilakukan perhitungan seluruh dataset dengan menggunakan *information gain*. Proses ini tentunya akan mendapatkan hasil berupa bobot untuk setiap atribut dalam dataset tersebut. Selanjutnya atribut yang nilainya kecil atau dalam arti tidak memiliki pengaruh besar nantinya dapat dihilangkan dari dataset untuk memperbaiki hasil klasifikasi.

2.3 Evaluasi Hasil

Setelah semua bobot atribut diketahui dengan menggunakan seleksi fitur *information gain* maka langkah berikutnya adalah evaluasi yaitu dengan membandingkan hasil klasifikasi dengan dataset asli (16 atribut) dan dengan menggunakan dataset yang telah dipilih. Evaluasi ini dilakukan hanya untuk memastikan bahwa usulan atribut yang digunakan atau pengurangan atribut benar benar berpengaruh terhadap hasil klasifikasi. Dan dengan demikian pula nantinya dapat disimpulkan bahwa atribut yang dihilangkan tidak terlalu berpengaruh terhadap klasifikasi jika hasil klasifikasi dengan pengurangan atribut terbukti lebih baik.

3. Hasil dan Pembahasan

3.1 Seleksi fitur *information gain*

Proses seleksi fitur sebenarnya dapat dilakukan dengan berbagai macam cara. Model untuk melakukan proses ini juga dapat dilakukan dengan tipe wrapper serta filter. *Information gain* merupakan salah satu metode filter yang banyak digunakan peneliti untuk proses seleksi fitur (Deng and Runger 2012). Proses ini dimulai dengan cara menghitung nilai entrophy dari seluruh atribut yang ada didalam klasifikasi. Berikut adalah proses perhitungan manual dari seleksi fitur *information gain* untuk klasifikasi persetujuan kredit.

Untuk menghitung nilai *information gain* setiap atribut dalam dataset persetujuan kredit dihitung terlebih dahulu nilai informasi harapan untuk kelas kredit lancar dan kredit macet. Dalam hal ini terdapat dua kelas yaitu lancar dan macet yang artinya nilai m dalam dataset ini adalah 2 yaitu (C_1 =Lancar, dan C_2 =Macet). Sedangkan jumlah record dalam dataset ini adalah 100 sehingga nilai $s=100$ dengan rincian 59 orang

macet dan 41 lainnya lancar dalam pembayaran kredit. Artinya dapat disimpulkan nilai $s_1=41$ dan nilai $s_2=59$. Setelah itu nilai informasi harapan untuk mengelompokkan kelas tersebut dapat dihitung dengan cara:

$$I(s_1, s_2) = -\frac{41}{100} \log \frac{41}{100} - \frac{59}{100} \log \frac{59}{100} = 0,9765$$

Misalkan atribut yang akan dihitung adalah jenis kelamin yang hanya memiliki 2 nilai berbeda yaitu Laki-laki dan Perempuan. Maka nilai v yang ada adalah $a_1=Laki-laki$ dan $a_2=Perempuan$. Selanjutnya atribut a tersebut dapat digunakan untuk memisahkan S kedalam 2 subset yang terinci dalam tabel 2 berikut

Tabel 2 Pemisahan nilai S

S	Jenis Kelamin	Lancar (s_1)	Macet (s_2)
S1	Laki-laki	21	23
S2	Perempuan	20	36

Sehingga $S=S_{11}+S_{12}+S_{21}+S_{22}=100$ sampel. Jika S_j merupakan jumlah sampel pada masing masing subset s_j maka informasi harapan dari subset Laki-laki Lancar dan Laki-laki Macet adalah sebagai berikut:

$$I(s_{11}, s_{21}) = -\frac{s_{11}}{s_1} \log \frac{s_{11}}{s_1} - \frac{s_{21}}{s_1} \log \frac{s_{21}}{s_1} = -\frac{21}{44} \log \frac{21}{44} - \frac{23}{44} \log \frac{23}{44} = 0,99851$$

Sedangkan informasi harapan dari subset Perempuan Lancar dan Perempuan Macet adalah sebagai berikut:

$$I(s_{12}, s_{22}) = -\frac{s_{12}}{s_2} \log \frac{s_{12}}{s_2} - \frac{s_{22}}{s_2} \log \frac{s_{22}}{s_2} = -\frac{20}{56} \log \frac{20}{56} - \frac{36}{56} \log \frac{36}{56} = 0,94029$$

Kemudian nilai *entropy* dari atribut Jenis kelamin dapat dihitung dari informasi harapan berdasarkan pemisahan kedalam subset berikut:

$$\begin{aligned} E(A) &= \frac{s_{11} + s_{21}}{S} I(s_{11}, s_{21}) + \frac{s_{12} + s_{22}}{S} I(s_{12}, s_{22}) \\ &= \frac{21 + 23}{44} 0,99851 + \frac{20 + 36}{56} 0,94029 = 0,97002 \end{aligned}$$

Dengan demikian nilai *information gain* dari atribut Jenis kelamin adalah:

$$\begin{aligned} \text{Gain (JnKel)} &= |I(S_1, S_2) - E(JnKel)| \\ &= |0,9765 - 0,97002| \\ &= \mathbf{0,0065} \end{aligned}$$

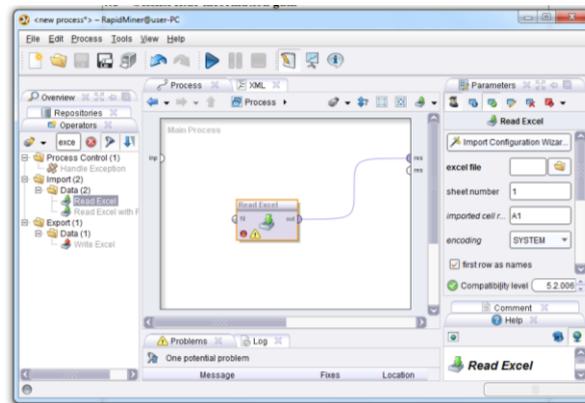
Kemudian perlakuan yang sama digunakan untuk menghitung keseluruhan atribut yang ada. Sehingga didapatkan hasil perhitungan seperti pada sub bab berikut.

3.2 Hasil seleksi fitur dengan tools rapid miner

Untuk proses seleksi fitur dengan menggunakan tools rapid miner dapat dilakukan dengan beberapa tahapan, tahapan berikut secara rinci akan dijelaskan dalam sub bab

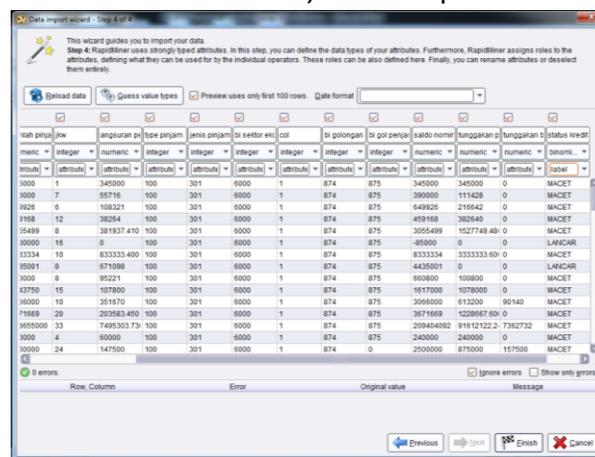
berikut. Pertama jalankan rapid miner kemudian import data file excell yang ada. File excell dipilih karena dataset yang ada bertipe file xls.

Gambar 1 Import data excell



Langkah berikutnya import konfigurasi dari dataset yang kita punya dengan memasukkan atribut beserta tipenya dan pemilihan label atau atribut tujuan yaitu status kredit. Gambar 2 merupakan pemilihan tipe atribut serta label dengan menggunakan rapid miner.

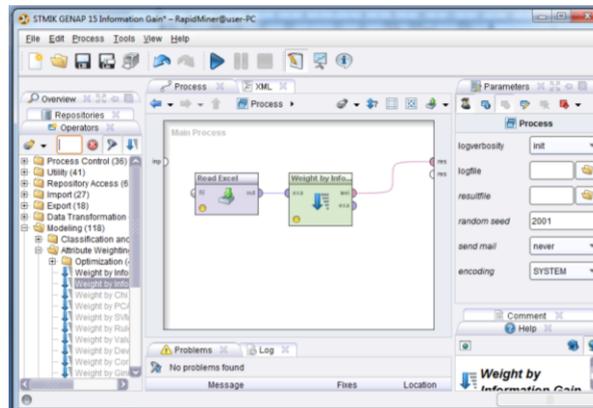
Gambar 2 Pemilihan jenis dan tipe atribut



Berikutnya setelah dataset berhasil diimport dalam program gunakan information gain untuk menghitung bobot dari keseluruhan atribut yang ada. Kita dapat mengambil information gain dengan cara drag and drop algoritma dari operator modelling dan atribut weighting kemudian pilih information gain. cara lain juga dapat kita gunakan yaitu dengan mencari menggunakan tombol search pada bagian atas operator. Gambar 3 adalah susunan algoritma information gain untuk menghitung bobot keseluruhan atribut data persetujuan kredit.

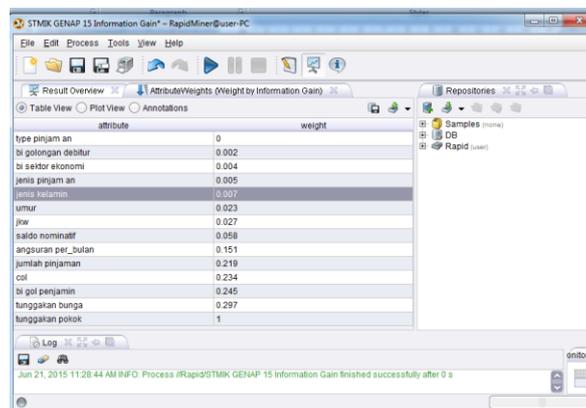
Gambar 3 Pemilihan jenis dan tipe atribut

Gambar 4 Susunan algoritma



Setelah tersusun sempurna maka ketika dijalankan akan muncul hasil seperti gambar 4 berikut.

Gambar 5 Hasil seleksi fitur



Berikutnya tabel 2 merupakan hasil perhitungan nilai *information gain* seluruh dataset persetujuan kredit. Terlihat bahwa tipe pinjaman mempunyai bobot 0 dalam perhitungan ini artinya atribut tipe pinjaman tidak berpengaruh terhadap klasifikasi. Sedangkan jenis kelamin mempunyai bobot 0,007 artinya sesuai dengan perhitungan manual. Dan nilai gain tertinggi yaitu atribut tunggakan pokok yang memiliki nilai gain sebesar 1.

Tabel 1 Hasil perhitungan *information gain*

Atribut	Information gain
type pinjam an	0.000
bi golongan debitur	0.002
bi sektor ekonomi	0.004
jenis pinjam an	0.005
jenis kelamin	0.007
umur	0.023
jkw	0.027

saldo nominatif	0.058
angsuran per_bulan	0.151
jumlah pinjaman	0.219
col	0.234
bi gol penjamin	0.245
tunggakan bunga	0.297
tunggakan pokok	1.000

4. Kesimpulan dan Saran

4.1 Kesimpulan

Dari hasil penelitian dapat diambil kesimpulan bahwa *information gain* dapat digunakan untuk mengetahui pengaruh atribut dataset terhadap klasifikasi. Atribut yang memiliki pengaruh terbesar adalah tunggakan pokok yaitu dengan bobot 1. Sedangkan atribut yang tidak memiliki pengaruh sama sekali dalam klasifikasi persetujuan kredit adalah tipe pinjaman dengan bobot *information gain* sebesar 0.

4.2 Saran

Dalam penelitian ini masih terdapat beberapa kekurangan antara lain perhitungan yang dilakukan menggunakan software bantu rapid miner. Dalam penelitian berikutnya dapat dibuat aplikasi bantu dengan menggunakan bahasa java ataupun php.

5. Referensi

- Alkaromi, M Adib. 2014. "Information Gain Untuk Pemilihan Fitur Pada Klasifikasi Heregistrasi Calon Mahasiswa Dengan Menggunakan K-NN."
- Alpaydin, Ethem. 2010. *Introduction to Machine Learning Second Edition*.
- Azhagusundari, B, and Antony Selvadoss Thanamani. 2013. "Feature Selection Based on Information Gain," no. 2: 18–21.
- Azwany, Faraby. 2010. "Sistem Pendukung Keputusan Pemberian Kredit Usaha Rakyat Pada Bank Syariah Mandiri Cabang Medan Menggunakan Metode AHP." *Program Studi Ilmu Komputer Departemen Ilmu Komputer Fakultas Matematika Dan Ilmu Pengetahuan Alam Universitas Sumatera Utara Medan*.
- Deng, Houtao, and George Runger. 2012. "Feature Selection via Regularized Trees," January. <http://arxiv.org/abs/1201.1587v3>.
- Dwi Cahyani, Bangun. 2013. "Sistem Pendukung Keputusan Persetujuan Permohonan Kredit Pinjaman Pada Bank Muamalat Indonesia Cabang Yogyakarta." *Sekolah Tinggi Manajemen Informatika Dan Komputer AMIKOM Yogyakarta*.
- Gallager, Robert G, and Life Fellow. 2001. "Claude E . Shannon : A Retrospective on His Life , Work , and Impact" 47 (7): 2681–95.

- Ginanjar Maburur, Angga, and Riani Lubis. 2012. "Penerapan Data Mining Untuk Memprediksi Kriteria Nasabah Kredit." *Jurnal Komputer Dan Informatika (KOMPUTA)* 1.
- Han, Jiawei, and Micheline Kamber. 2006. *Data Mining: Concepts and Techniques Second Edition*. Elsevier. Elsevier.
- Ian H Witten. Eibe Frank. Mark A Hall. 2011. *Data Mining 3rd*.
- Koprinska, Irena. 2010. "Feature Selection for Brain-Computer Interfaces," 100–111.
- Larose, Daniel T. 2005. *Discovering Knowledge in Data: An Introduction to Data Mining*. John Wiley & Sons.
- Maimoon, Oded, and Lior Rokach. 2010. *Data Mining and Knowledge Discovery Handbook*. Vol. 40. doi:10.1002/1521-3773(20010316)40:6<9823::AID-ANIE9823>3.3.CO;2-C.
- Melissa, Ira, and Raymond S Oetama. 2013. "Analisis Data Pembayaran Kredit Nasabah Bank Menggunakan Metode Data Mining" IV (1): 18–27.
- Novakovic, Jasmina. 2010. "The Impact of Feature Selection on the Accuracy of 1DwYH Bayes Classifier" 2: 1113–16.
- Prasetyo, Eko. 2012. *Data Mining Konsep Dan Aplikasi Menggunakan Matlab*. Yogyakarta: Andi Offset.
- Santosa, Budi. 2007. *Data Mining Teknik Pemanfaatan Data Untuk Keperluan Bisnis*. Edisi Pert. Yogyakarta: Graha Ilmu.
- Suhari, Yohanes, Muji Sukur, and Sri Eniyati. 2009. "SISTEM PENDUKUNG KEPUTUSAN PEMBERIAN KREDIT PADA PT . BPR ARTAMANUNGGAL ABADI MRANGGEN." *Dinamika Informatika Fakultas Teknologi Informasi Universitas Stikubank Semarang I* (1).
- W, Yogi Yusuf, F Rian Pratikto, and A S Vivianne. 2009. "EVALUASI PEMOHON KREDIT MOBIL DI PT ' X ' DENGAN MENGGUNAKAN TEKNIK DATA MINING DECISION TREE." *Simposium Nasional RAPI VIII*, 42–49.
- Wachid K, Achmad. "Desain Model Sistem Pendukung Keputusan Kelayakan Kredit Mobil."
- Wahyu Oktaputra, Alif, and Edi Noersasongko. 2014. "Sistem Pendukung Keputusan Kelayakan Pemberian Kredit Motor Menggunakan Metode Simple Additive Weighting Pada Perusahaan Leasing HD Finance." *Program Studi Sistem Informasi - S1, Fakultas Ilmu Komputer Universitas Dian Nuswantoro, Semarang*, 1–9.
- Wasana, Emil. 2010. "Sistem Pendukung Keputusan Kelayakan Pengajuan Kredit Motor Menggunakan Metode Scoring System." *Jurusan Siste Informasi, Sekolah Tinggi Manajemen Informatika & Teknik Komputer Surabaya*, 1–10.
- Witten, Ian H, Eibe Frank, and Mark A. Hall. 2011. *Data Mining: Practical Machine Learning Tools and Techniques 3rd Edition*. Elsevier.
- Zahid, Ahmad. 2013. "Sistem Pendukung Keputusan Persetujuan Penerimaan Pinjaman Di PD.BPR BKK Lasem Dengan Menggunakan Metode Bayes." *Sekolah Tinggi Manajemen Informatika Dan Komputer AMIKOM Yogyakarta*.