



Artikel

Pendugaan Kredit Macet Pada Koperasi Simpan Pinjam Flamboyan Binaan PPSW Jakarta Dengan Menggunakan Komparasi Algoritma *Naïve Bayes* dan *C4.5*

Renaldi¹, Yusuf Kurnia², Ellysha Dwiyanthi Kusuma³¹Teknologi Indonesia, Universitas Buddhi Dharma, Banten, Indonesia

SUBMISSION TRACK

Recieved:
Final Revision:
Available Online:

KEYWORD

Naïve Bayes, *C4.5*, Komparasi, Pendugaan

KORESPONDENSI

Phone: 085885907575
E-mail: lim.renaldi123@gmail.com

A B S T R A C T

Data mining sering kali digunakan dalam bidang keuangan salah satunya adalah koperasi. Menurut *UU no 25 tahun 1992*, yang dimaksud dengan Koperasi adalah badan usaha yang beranggotakan orang seorang atau badan hukum koperasi dengan melandaskan kegiatan berdasarkan prinsip koperasi sekaligus sebagai gerakan ekonomi rakyat yang berdasarkan pada asas kekeluargaan. Peraturan Menteri koperasi dan Usaha Kecil dan Menengah (*Permen Kop & UMKM*) Nomor *15/Per/M.KUKM/IX/2015* yang menyebutkan bahwa modal sendiri *KSP* itu terdiri dari simpanan pokok, simpanan wajib, cadangan yang disisihkan dari sisa hasil usaha, hibah dan simpanan lain yang memiliki karakteristik sama dengan simpanan wajib. Salah satu hal yang perlu diperhatikan yaitu mengenai pemberian kredit atau peminjaman dalam koperasi Flamboyan, yang dalam penelitian ini banyak terjadi pengkreditan macet yang terjadi di koperasi Flamboyan. Dengan memakai teknik Data Mining data yang banyak tersebut dapat dimanfaatkan secara optimal. Dari permasalahan diatas dapat diatasi dengan memanfaatkan Teknik data *mining* yaitu Pendugaan Kredit Macet Pada Koperasi Simpan Pinjam Flamboyan Binaan PPSW Jakarta Dengan Memakai Komparasi Algoritma *Naive Bayes* dan *C4.5*. Algoritma yang dipakai ke dalam sistem adalah hasil terbaik dari komparasi *Naive Bayes* dan *C4.5* berdasarkan dari data koperasi Flamboyan. Hasil yang didapatkan dari pengolahan data komparasi antara algoritma *Naive Bayes* dan *C4.5* dengan memakai dataset yang berjumlah 2282 data transaksi didapat hasil akurasi dari algoritma *Naive Bayes* sebesar 69.19% dan algoritma *C4.5* sebesar 71.87%, berdasarkan dari hasil akurasi menyatakan bahwa algoritma *C4.5* lebih unggul dari algoritma *Naive Bayes*. Maka hasil dari pohon keputusan *C4.5* diterjemahkan kedalam sistem pendugaan kredit macet pada koperasi Flamboyan.

PENGANTAR

Banyak definisi bagi istilah *Data Mining* dan belum ada yang dibakukan atau disepakati semua pihak. Namun demikian, istilah ini memiliki hakikat (*Notion*) sebagai disiplin ilmu yang tujuan utamanya adalah untuk menemukan, menggali, atau menambangkan pengetahuan dari data atau informasi yang kita miliki. Kegiatan inilah yang menjadi garapan atau perhatian utama dari disiplin ilmu *data mining* (Sani Susanto, 2010:2). *Data Mining* menurut David Hand, Heikki Mannila dan Padhraic dari MIT adalah Analisa terhadap data yang biasanya berukuran besar untuk menemukan hubungan yang jelas serta menyimpulkan yang belum diketahui sebelumnya dengan cara terkini dipahami dan berguna bagi pemilik data tersebut.

Data mining sering kali digunakan dalam bidang keuangan salah satunya adalah koperasi. Menurut *UU no 25 tahun 1992*, yang dimaksud dengan Koperasi adalah badan usaha yang beranggotakan orang seorang atau badan hukum koperasi dengan melandaskan kegiatan berdasarkan prinsip koperasi sekaligus sebagai gerakan ekonomi rakyat yang berdasarkan pada asas kekeluargaan. Peraturan Menteri koperasi dan Usaha Kecil dan Menengah (*Permen Kop & UMKM*) Nomor 15/Per/M.KUKM/IX/2015 yang menyebutkan bahwa modal sendiri *KSP* itu terdiri dari simpanan pokok, simpanan wajib, cadangan yang disisihkan dari sisa hasil usaha, hibah dan simpanan lain yang memiliki karakteristik sama dengan simpanan wajib. Koperasi simpan pinjam merupakan koperasi yang bergerak dalam lapangan usaha pembentukan modal melalui tabungan – tabungan para anggota secara teratur dan terus – menerus untuk selanjutnya dipinjamkan kepada untuk anggota secara cepat, biaya murah, dimudahkan dan tepat untuk tujuan produktif dan untuk kesejahteraan Ninik Widiyanti dan Sunindhia, (2009:198). Sebagai upaya koperasi dalam mendukung pelayanan yang baik diperlukan tata kerja yang tertib, rapih dan teliti sehingga akan menghasilkan informasi yang cepat, akurat

dan tepat waktu sesuai kebutuhan. Sedangkan dalam koperasi simpan pinjam tentunya banyak data yang dari setiap tahunnya bertambah, sehingga dari data tersebut tidak dapat diolah secara teratur dan nantinya akan hanya menjadi arsip saja.

Salah satu hal yang perlu diperhatikan yaitu mengenai pemberian kredit atau peminjaman dalam koperasi Flamboyan, yang dalam penelitian ini banyak terjadi pengkreditan macet yang terjadi di koperasi Flamboyan. Dengan memakai teknik *Data Mining* data yang banyak tersebut dapat dimanfaatkan secara optimal. Dengan memanfaatkan data – data tersebut dapat membantu mengatasi kesalahan dalam pemberian pinjaman yang sering terjadi pada koperasi simpan pinjam Flamboyan ini. Dengan memakai cara data *mining* ini dapat membantu pihak koperasi Flamboyan dalam mengatasi pengkreditan macet peminjaman menjadi lebih lancar dan semakin teratur. Selain itu juga dapat menjadi sarana pengambilan keputusan dalam pemberian kredit peminjaman terhadap anggota koperasi Flamboyan.

Dari permasalahan diatas dapat diatasi dengan memanfaatkan Teknik data *mining* yaitu Pendugaan Kredit Macet Pada Koperasi Simpan Pinjam Flamboyan Binaan PPSW Jakarta Dengan Memakai Komparasi Algoritma *Naive Bayes* dan *C4.5*. Dengan harapan setelah diolah dengan data *mining* dapat membantu menemukan informasi dalam pendugaan pengkreditan macet peminjaman modal pada koperasi Flamboyan.

I. METODE

Dalam penelitian ini penulis menerapkan Metode Klasifikasi dengan algoritma *C4.5* dalam aplikasi yang dapat memprediksi kredit macet yang akan terjadi.

1.1. Metode Klasifikasi

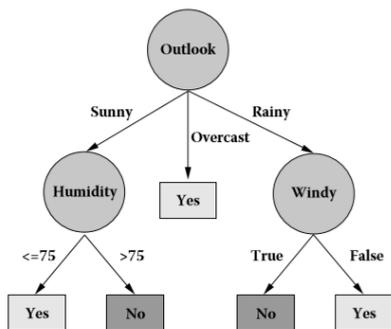
[1] Dalam *Data Mining* ada yang disebut pengklasifikasian, klasifikasi pertama kali diterapkan pada bidang tanaman yang mengklasifikasikan suatu spesies tertentu,

seperti yang dilakukan oleh Carolus von Linne (atau yang dikenal dengan nama Carolus Linnaeus) yang pertama kali mengklasifikasi spesies berdasarkan karakteristik fisik. Klasifikasi merupakan suatu pekerjaan menilai objek data untuk memasukkannya ke dalam kelas tertentu dari sejumlah kelas yang tersedia. Dalam klasifikasi ada dua pekerjaan utama yang dilakukan yaitu (1) pembangunan model sebagai prototipe untuk disimpan sebagai memori dan penggunaan model tersebut untuk melakukan pengenalan / klasifikasi / prediksi pada suatu objek data lain agar diketahui di kelas mana objek data tersebut dalam model yang sudah disimpannya.

1.2. C4.5 (Decision Tree)

C4.5 adalah bagian dari algoritma untuk klasifikasi dalam pembelajaran *machine learning* dan *data mining*. C4.5 merupakan algoritma yang cocok digunakan untuk masalah klasifikasi pada *machine learning* dan *data mining*.

Semua metode penginduksian pohon mulai dengan booting yang tidak menghadirkan perangkat, given dataset dan secara terpecah membagi data menjadi himpunan bagian yang lebih kecil dengan menguji atribut yang diberikan pada setiap node. Subpohon menunjukkan partisi dari dataset asli yang memenuhi nilai atribut tertentu yang ditentukan. Proses ini secara tipikal melanjutkan hingga subset adalah "murni," yaitu, semua contoh dalam subset jatuh di kelas yang sama, di mana saat pohon tumbuh diakhiri.



Gambar 2. 1 Pohon keputusan

Gambar 2.2 menyajikan dataset "golf" klasik, yang dibundel dengan instalasi C4.5. Seperti yang dinyatakan sebelumnya, tujuannya adalah untuk memprediksi apakah kondisi cuaca pada satu paket dapat digunakan untuk bermain serigala. Ingat bahwa sebagian dari fitur ini dinilai terus menerus sementara yang lain bersifat kategori. Gambar 2.2 mengilustrasikan pohon yang diinduksi oleh C4.5 sebagai data pelatihan dan opsi kesalahan [2].

1.3. Data Set

Bukan data mining Namanya jika tidak ada set data yang diolah didalamnya. Kata 'data' dalam terminology statistic adalah kumpulan objek dengan atribut – atribut tertentu, di mana objek tersebut adalah individu berupa data di mana setiap data sejumlah atribut. Atribut tersebut berpengaruh pada dimensi dari data, semakin banyak atribut/fitur maka semakin besar dimensi data. Kumpulan data – data membentuk set data. Dalam buku ini kadang menyebut data, kadang menyebut vector, keduanya mempunyai maksud yang sama.

Berikut tiga jenis set data yang dikenal dan masing – masing penggolongannya:

- 1.Record
 - a.Matriks data
 - b.Data transaksi
 - c.Data dokumen
- 2.Graph
 - a. Word Wide Web (WWW)
 - b.Struktur molekul
- 3.Ordered data set
 - a.Data spasial
 - b.Data temporal
 - c.Data sekuensial
 - d.Data urutan genetic (*genetic sequence*)

Set data yang berisi kumpulan data, dengan semua data mempunyai sejumlah atribut (fitur) numerik yang sama, dapat dipandang sebagai vector (data) dalam wilayah multidimensi, dimana setiap dimensi (fitur) memrepresentasikan atribut berbeda yang menggambarkan objek/data. Data matriks merupakan jenis record data yang paling umum dan banyak digunakan pada aplikasi – aplikasi statistic [1].

1.4. Data Mining

“Mendefinisikan *data mining* sebagai proses untuk mendapatkan informasi yang berguna dari Gudang basis data yang besar. Data yang sedang kita bicarakan dalam konteks *data mining* yang sebenarnya adalah data yang di kelola dalam suatu *database container*” [2].
 “Istilah *Data Mining* memiliki beberapa padanan, seperti *knowledge siccovery* ataupun *pattern recognition*. Kedua istilah tersebut sebenarnya memiliki ketepatannya masing – masing. Istilah *knowledge discovery* atau penemuan pengetahuan tepat digunakan karena tujuan utama dari data mining memang untuk mendapatkan pengetahuan yang masih tersembunyi didalam bongkahan data” [3].

II. HASIL

2.1. Model Rapid Miner

Dari hasil komparasi antara algoritma *Naïve Bayes* dan C4.5 dengan memakai dataset yang berjumlah 2282 data transaksi didapat hasil akurasi dari algoritma *Naïve Bayes* sebesar 69.19% dan algoritma C4.5 sebesar 71.87%. berdasarkan dari hasil akurasi menyatakan bahwa algoritma C4.5 lebih unggul dari algoritma *Naïve Bayes*.

	true lancar	true macet
pred. lancar	1612	586
pred. macet	56	28
class recall	96.64%	4.56%

Gambar 3. 1 Hasil akurasi cross validation Pertama C4.5 (Decision tree Tree)

	true lancar	true macet
pred. lancar	1320	355
pred. macet	348	259
class recall	79.14%	42.18%

Gambar 3. 2 Hasil akurasi Cross Validation Pertama Naïve Bayes

2.2. Algoritma C4.5

Dari hasil komparasi antara C4.5 dan *Naive Bayes* akurasi terbaik dimenangkan oleh

algoritma C4.5, oleh sebab itu penulis memakai algoritma terbaik kedalam perhitungan manual dan sistem yang dibuat. C4.5 merupakan bagian dari algoritma untuk klasifikasi dalam pembelajaran *machine learning* dan data mining. Algoritma C4.5 memakai konsep *information gain* dan *entropy reduction* untuk memilih pembagian yang optimal dan menghasilkan pohon keputusan. Tahapan dalam membuat sebuah pohon keputusan dengan algoritma C4.5 adalah sebagai berikut :

1. Mempersiapkan data

Dalam penelitian ini penulis memakai data training sebanyak 60 data yang memiliki semua kondisi klasifikasi yang diambil secara acak dari data set.

2. Menghitung Entropy

Menentukan akar dari pohon dengan menghitung nilai *gain* yang tertinggi dari masing – masing atribut berdasarkan nilai *index entropy* terendah. Sebelumnya terlebih dahulu dihitung nilai *index entropy* dengan rumus :

$$\text{Entropi (S)} = \sum_{j=1}^k - p_i \cdot \log_2 \cdot p_i$$

Keterangan :

S : himpunan kasus

k : jumlah partisi i

pi : probabilitas yang di dapat dari

Sum

3. Menghitung Gain

Hitung gain dengan rumus :

$$\text{Gain (S, A)} = \text{Entropy (S)}$$

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

Untuk mendapatkan pohon keputusan dari data yang digunakan maka harus dilakukan perhitungan dengan mengimplementasikan rumus perhitungan Entropy dan Gain yang ada diatas yaitu sebagai berikut :

Pada tahap ini merupakan perhitungan untuk menentukan kondisi atribut mana yang pertama kali menentukan atribut mana yang memiliki gain paling tinggi. Hasil Node 1 ini menunjukkan bahwa gain yang paling besar terjadi pada atribut pinjaman, maka pinjaman

merupakan daun pertama dalam pohon keputusan.

1. Perhitungan Entropy Total

Entropy (Total)

$$= \left(-\frac{49}{60} * \log_2 \left(\frac{49}{60} \right) \right) + \left(-\frac{11}{60} * \log_2 \left(\frac{11}{60} \right) \right) = 0.687315093$$

2. Perhitungan Entropy atribut Status Kerja

Entropy (Status Kerja, Y)

$$= \left(-\frac{26}{33} * \log_2 \left(\frac{26}{33} \right) \right) + \left(-\frac{7}{33} * \log_2 \left(\frac{7}{33} \right) \right) = 0.745517843$$

Entropy (Status Kerja, N)

$$= \left(-\frac{5}{5} * \log_2 \left(\frac{5}{5} \right) \right) + \left(-\frac{0}{5} * \log_2 \left(\frac{0}{5} \right) \right) = 0$$

Entropy (Status Kerja, T)

$$= \left(-\frac{18}{22} * \log_2 \left(\frac{18}{22} \right) \right) + \left(-\frac{4}{22} * \log_2 \left(\frac{4}{22} \right) \right) = 0.684038436$$

• Perhitungan Gain Status Kerja

Gain (Status Kerja)

$$= 0.687315093 - \left(\left(\frac{33}{60} * 0.746 \right) + \left(\frac{5}{60} * 0 \right) + \left(\frac{22}{60} * 0.684 \right) \right) = 0.026466186$$

Dari hasil contoh perhitungan node diatas maka terciptalah hasil pohon keputusan seperti berikut ini :

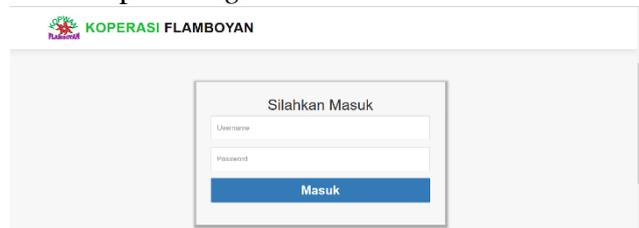


Gambar 4. 1 Hasil Pohon Keputusan Rapid Miner

III. PEMBAHASAN

3.1. Tampilan Program

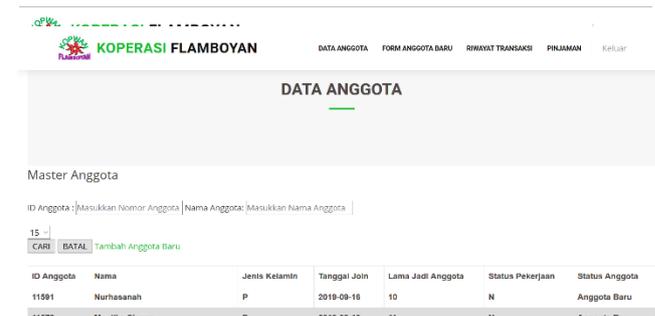
a. Tampilan Login



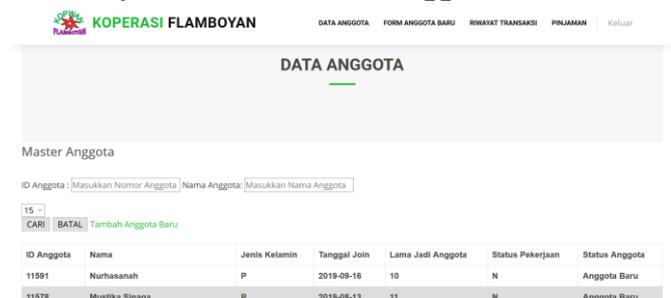
Gambar 3.1 Halaman Login

b. Tampilan Menu Utama

Gambar 3.2 Halaman Utama



c. Tampilan Halaman Data Anggota



Gambar 3.3 Halaman Data Anggota

pengembangan lebih lanjut adalah sebagai berikut :

1. Dalam pengembangan lebih lanjut disarankan agar *logic* pada sistem pendugaan memakai perhitungan algoritma C4.5 yang di aplikasikan langsung kedalam sistem.
2. Pada sistem pendugaan kredit macet untuk pengembangan lebih lanjutnya agar ditambahkan fitur menu laporan, agar dalam sistem ini memiliki fitur laporan transaksi yang tersimpan didalam sistem.
3. Untuk pengembangan selanjutnya disarankan untuk dapat menambah atribut dan variable data agar mendapat hasil perhitungan yang lebih baik.

DAFTAR PUSTAKA

- [1]. Prasetyo, Eko. 2012. *Data Mining Konsep Dan Aplikasi Menggunakan Matlab*. Yogyakarta: Andi
- [2]. Xindong Wu dan Vipin Kumar. 2009. *The Top Ten Algorithms in Data Mining*. Minnesota USA : Chapman & Hall/CRC.
- [3]. Finn Lee S dan Juan Santana. 2010. *Data Mining : Meramalkan Bisnis Perusahaan*. Jakarta: PT. Elex Media Komputindo.

BIOGRAFI

Renaldi, S.Kom, dilahirkan di Tangerang, 10 maret 1998. Sekolah Dasar dilaksanakan di SD Dharma Widya, Kota Tangerang, SMP Dharma Widya, Kota Tangerang, SMK Buddhi, Kota Tangerang. Pendidikan S-1 (Teknik Informatika – Universitas Buddhi Dharma) lulus tahun 2020.

Yusuf Kurnia, Saat ini bekerja sebagai dosen Tetap pada Program Studi Teknik Informatika di Universitas Buddhi Dharma.

Ellysha Dwiyanthi Kusuma, Saat ini bekerja sebagai dosen Tetap pada Program Studi Teknik Informatika di Universitas Buddhi Dharma.